

RÉGULATION DE L'EXPRESSION GÉNÉTIQUE PAR LE
 MG^{2+} :
CAS DU RIBORÉGULATEUR *MGTA* ET
IDENTIFICATION DE NOUVEAUX GÈNES RÉGULÉS
PAR LE MG^{2+}

par

Mélanie Geffroy

thèse présentée au Département de biologie en vue

de l'obtention du grade de docteur ès Science (Ph.D.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, août 2018

Le 15 août 2018,

le jury a accepté le mémoire de Madame Mélanie Geffroy dans sa version finale.

Membres du jury

Professeur Daniel Lafontaine

Directeur de recherche

Département de biologie

Professeur Eric Massé

Codirecteur de recherche

Département de biochimie

Professeur Pascal Chartrand

Évaluateur externe

Département de biochimie, Université de Montréal

Professeur Sébastien Rodrigue

Évaluateur interne

Département de biologie

Professeur Vincent Burrus

Président-rapporteur

Département de biologie

SOMMAIRE

Les bactéries sont directement exposées à un environnement qui peut changer radicalement. Afin de s'adapter à de telles variations, elles doivent être capables de réguler rapidement et finement l'expression génique. La concentration en ion magnésium (Mg^{2+}) par exemple peut varier beaucoup d'un environnement à un autre. Cet ion étant responsable de la stabilité et le fonctionnement de nombreuses molécules et complexes cellulaires, il est important pour la bactérie de pouvoir contrôler sa concentration intracellulaire, quelle que soit la disponibilité dans le milieu. L'étude de la régulation de l'expression génique par le Mg^{2+} est pour ces raisons le sujet général de mon doctorat.

Les riborégulateurs sont des régions d'ARN structurées, retrouvées principalement dans les régions 5' des ARNm bactériens, capables de reconnaître et lier spécifiquement un ligand afin de contrôler l'expression génique. La liaison du ligand est responsable d'un changement de conformation de l'ARN qui permet la régulation par divers mécanismes qui peuvent agir aussi bien au niveau de la transcription, de la traduction, que de la dégradation de l'ARN. Au cours de mes travaux, j'ai étudié le riborégulateur *mgtA* chez *E. coli*. Le ligand de ce riborégulateur est l'ion Mg^{2+} et le gène régulé code pour le transporteur de Mg^{2+} MgtA. Les résultats de mon étude ont permis de déterminer que ce riborégulateur utilise un mécanisme de régulation de la transcription et que cette régulation est dépendante de la traduction d'un peptide leader, mécanisme unique à ce jour.

Dans le cadre de l'étude du riborégulateur, j'ai effectué une expérience de *profilage ribosomique* qui a généré de nombreuses données que j'ai par la suite analysées. Les résultats

obtenus m'ont permis de montrer que la régulation par le Mg^{2+} est plus étendue que prévu par la littérature actuelle.

Mots-clefs : Riborégulateur, *mgtA*, Mg^{2+} , ARN, peptide leader, *profilage ribosomique*

REMERCIEMENTS

Je tiens tout d'abord à remercier mon directeur de thèse, le Pr Daniel Lafontaine, de m'avoir accueillie au sein de son laboratoire, de m'avoir fait confiance, de m'avoir soutenue tout au long de mon doctorat. Merci de m'avoir encouragée, pour tes bons conseils et ta rigueur scientifique. Merci également à mon codirecteur, le Pr Eric Massé. Merci, Eric, d'avoir accepté de me soutenir dans ce projet. Nos discussions et les membres de ton laboratoire m'ont beaucoup apporté. J'ai beaucoup appris durant ces quelques années passées dans vos deux laboratoires!

Un grand merci à mes conseillers, le Pr Vincent Burrus et le Pr Sébastien Rodrigue, nos réunions ont été enrichissantes et motivantes. Vos précieux conseils et votre expertise ont contribué à la réalisation de ce projet. Je tiens à exprimer ma sincère gratitude au Pr Pascal Chartrand pour avoir accepté d'être examinateur externe de mon jury de thèse. Je remercie également nos collaborateurs, le Pr Joseph T. Wade et son étudiante Anne Stringer.

Un grand merci à tous les membres des laboratoires Lafontaine et Massé que j'ai pu côtoyer pendant toutes ces années! Un grand merci à Karine qui a beaucoup contribué à cette thèse ; à Laurène pour son aide précieuse et pour son amitié ; à Adrien pour sa capacité à nous inciter à nous surpasser et pour nos discussions scientifiques ; à Anne-Marie, de prendre soin du laboratoire et sans qui nous sommes tous perdus ; à JF pour ses encouragements et sa précieuse amitié ; à Pierre, mon stagiaire éphémère et ami moins éphémère! ; à Alex pour son soutien inconditionnel et ses conseils ; à Anne-Sophie, de prendre la suite de mes travaux avec tant de motivation ; à Pat, Jo, Marie-Claude, David, Élodie, JC, Auréliane, Julie-Anna, Max, Antony et les autres. Merci également à Dom et Vince du laboratoire Rodrigue pour leur aide pour la partie bio-informatique!

Et surtout un merci à ma famille pour leur soutien inconditionnel et sans qui je ne serais pas là aujourd'hui. Et merci à Bruno de me soutenir malgré la fin de thèse!

TABLE DES MATIÈRES

Sommaire.....	iv
Remerciements	vi
Liste des abréviations	xi
Liste des tableaux	xiii
Liste des figures.....	xiv
CHAPITRE 1 Introduction	1
1.1. Mise en contexte	1
1.1.1. Importance de l'ion Mg^{2+}	1
1.1.2. Transporteurs de Mg^{2+}	3
1.2. Régulation de l'expression génique	4
1.2.1. Expression génique.....	4
1.2.1.1. Transcription	4
1.2.1.2. Traduction	8
1.2.1.3. Dégradation de l'ARNm	12
1.2.2. La régulation de l'expression génétique.....	17
1.2.2.1 Transcription.....	17
1.2.2.2. Dégradation de l'ARNm.....	19
1.2.2.3. Traduction.....	19
1.2.3. Les petits ARN	21
1.2.4. Les riborégulateurs.....	24
1.2.4.1. Riborégulateurs transcriptionnels	27
1.2.4.2. Riborégulateurs contrôlant la stabilité de l'ARNm	28
1.2.4.3. Riborégulateurs traductionnels	30

1.2.5. Les peptides leaders	31
1.2.6. La régulation par le Mg^{2+}	37
1.3. Le riborégulateur <i>mgtA</i> chez <i>E. coli</i>	39
1.3.1. Le riborégulateur chez <i>S. enterica</i>	39
1.3.2. Le riborégulateur <i>mgtA</i> chez <i>Yersinia enterocolitica</i>	42
1.3.3. Le riborégulateurs <i>mgtA</i> chez <i>E. coli</i>	42
1.4. Objectifs du projet de recherche	43
CHAPITRE 2 : Matériel et Méthodes	45
2.1. Souches, plasmides et oligos	45
2.2. SHAPE natif	54
2.3. Cartographie Enzymatique.....	54
2.4. Essais β -galactosidase.....	55
2.5. Transcriptions à cycle unique	55
2.6. Immunobuvardage de type northern	56
2.7. Dégradation <i>in vitro</i>	56
2.8. ChIP-qPCR	56
2.9. Identification des sites de pause.....	57
2.10. Profilage ribosomique.....	57
2.11 Traduction <i>in vitro</i>	58
2.12. Analyse des données de séquençage.....	58
CHAPITRE 3 Le riborégulateur <i>mgtA</i> chez <i>E. coli</i>	60
3.1. La région 5' non traduite de <i>mgtA</i> peut former deux structures alternatives.....	60
3.2. La région 5' non traduite régule l'expression de <i>mgtA</i> par le Mg^{2+} au niveau de l'ARNm	66
3.3. Le riborégulateur n'utilise pas de terminateur de transcription intrinsèque.	70
3.4. Le riborégulateur n'utilise pas un terminateur dépendant de Rho	71

3.5. Le dégradosome n'est pas impliqué dans la régulation	73
3.6. L'élongation de la transcription est affectée par le Mg^{2+}	77
CHAPITRE 4 Le peptide leader <i>mgtL</i>	82
4.1. Présence d'un cadre de lecture dans le riborégulateur	82
4.2. Le peptide est traduit et son expression est régulée par le Mg^{2+}	83
4.3. La traduction du peptide MgtL est nécessaire à l'expression de <i>mgtA</i>	85
4.4. La régulation de la traduction du peptide est suffisante pour réguler l'expression de <i>mgtA</i>	89
4.5. La séquence du peptide est importante pour le contrôle de l'expression de <i>mgtA</i>	93
4.6. Mécanisme de régulation du riborégulateur et du peptide leader	97
CHAPITRE 5 Analyse des données de profilage ribosomique	106
5.1. Le profilage ribosomique pour évaluer l'expression génique.....	106
5.2. Qualité des données	109
5.3. Expression différentielle des gènes régulés par le Mg^{2+}	117
5.5. Identification de nouveaux gènes du stimulon Mg^{2+}	120
5.5.1. Le gène <i>corA</i>	121
5.5.2. Le gène <i>ybeX</i>	126
5.5.3. Le gène <i>ynfB</i>	129
5.5.4. L'opéron <i>hisLGDC</i>	131
5.5.5. Le gène <i>ybjX</i>	132
5.5.6. L'opéron <i>gsiAB</i>	133
5.5.7. Le gène <i>mlaF</i>	136
5.5.8. Le gène <i>cysD</i>	136
5.5.9. Le gène <i>ompX</i>	138
CHAPITRE 6 Discussion et conclusion générale	140
6.1. Conclusion générale.....	140

6.2. Mg^{2+} et phase de croissance.....	141
6.3. Rôle du peptide leader	142
6.4. Comparaison avec <i>S. enterica</i>	144
6.5. Riborégulateur ou peptide leader?	145
6.6. Régulation par un pARN	147
6.7. Dégradation.....	147
6.8. Régulation de la traduction	148
6.9. Modèle proposé.....	149
6.10. Découverte de nouveaux gènes régulés par le Mg^{2+}	151
6.11. Contributions	152
Bibliographie	155

LISTE DES ABRÉVIATIONS

ADN	Acide désoxyribonucléique
Anti-SD	anti-Shine Dalgarno
ARN	Acide ribonucléique
ARNm	ARN messenger
ARNr	ARN ribosomique
ARNt	ARN de transfert
ARNt-aa	ARNt-aminoacylé
ATP	Adénosine triphosphate
ATPase	Adénosine triphosphatase
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
Da	Dalton
CI30S	complexe d'initiation 30S
di-AMPC	Diadénosine monophosphate cyclique
di-GMPc	Diguanosine monophosphate cyclique
<i>E. coli</i>	<i>Escherichia coli</i>
FMN	Flavine mononucléotide
FRET	Förster Resonance Energy Transfer
GlcN6P	Glucosamine-6-phosphate
Mg ²⁺	Ion magnésium
miRNA	Micro-ARN
NET-seq	<i>Nascent Elongating Transcript sequencing</i>
NTP	Nucléotide triphosphate
Opéron his	Opéron histidine

Opéron trp	Opéron tryptophane
Ops	<i>Operon polarity suppressor</i>
pARN	petit ARN
PdARN	polymérase d'ARN
PNPase	Polynucléotide phosphorylase
PréQ1	7-aminométhyl-7-déazaguanine
RBS	Ribosome Binding Site
RNase	Ribonucléase
<i>rut</i>	<i>Rho utilization</i>
<i>S. enterica</i>	<i>Salmonella enterica</i>
SAM	S-adénosyl-méthionine
SAH	S-adénosyl-monocystéine
SD	Shine-Dalgarno
SHAPE	<i>Selective 2'-Hydroxyl Acylation analyzed by Primer Extension</i>
siRNA	Petit ARN interférent
sm-FRET	single molecule FRET
TPP	Thiamine pyrophosphate
trp	tryptophane
TRAP	<i>Tryptophan Regulated Attenuation Protein</i>
UTR	<i>Untranslated Region</i>

LISTE DES TABLEAUX

Tableau 1.	Liste des riborégulateurs connus	25
Tableau 2.	Liste des souches	45
Tableau 3.	Tableau des amorces	48
Tableau 4.	Stratégies de PCR avec les amorces utilisées	51
Tableau 5.	Gènes régulés significativement par le Mg^{2+} d'après le profilage ribosomique	120

LISTE DES FIGURES

Figure 1.	Le cœur de la polymérase d'ARN	5
Figure 2.	Élongation de la traduction	10
Figure 3.	La dégradation des ARNm chez <i>E. coli</i>	14
Figure 4.	Mécanismes de régulation des pARN	23
Figure 5.	Mécanismes de régulation transcriptionnelle	28
Figure 6.	Mécanisme de régulation d'un riborégulateur dégradationnel	29
Figure 7.	Mécanisme de régulation d'un riborégulateur traductionnel	30
Figure 8.	Mécanisme de régulation de l'opéron <i>trp</i> chez <i>E. coli</i>	33
Figure 9.	Régulation par le système à deux composants PhoQ/PhoP	38
Figure 10.	Structures du riborégulateur <i>mgtA</i> chez <i>S. enterica</i>	40
Figure 11.	Prédictions des structures adoptées par le riborégulateur <i>mgtA</i>	61
Figure 12.	Alignement de structures de la région 5' de <i>mgtA</i>	62
Figure 13.	Cartographie de la structure du riborégulateur <i>mgtA</i> chez <i>E. coli</i>	65

Figure 14. Deux prédictions mfold pour la structure active	66
Figure 15. Régulation de l'expression de <i>mgtA</i>	69
Figure 16. Le terminateur canonique n'est pas responsable de la régulation	71
Figure 17. La protéine Rho n'est pas impliquée dans le mécanisme de régulation	73
Figure 18. Demi-vie de l'ARNm <i>mgtA</i>	74
Figure 19. Implication du dégradosome dans la régulation	76
Figure 20. ChIP-qPCR de la polymérase d'ARN en faible et fort Mg^{2+}	78
Figure 21. Les sites de pause de la polymérase d'ARN dans le riborégulateur <i>mgtA</i>	81
Figure 22. Une région du riborégulateur peut être traduite <i>in vivo</i>	83
Figure 23. La technique de profilage ribosomique révèle l'expression et la régulation de <i>mgtL</i>	85
Figure 24. La traduction du peptide est nécessaire à l'expression et la régulation de <i>mgtA</i>	87
Figure 25. La surexpression de <i>mgtL</i> en <i>trans</i> n'affecte pas l'expression de <i>mgtA</i>	88
Figure 26. Le contrôle de la traduction de <i>mgtL</i> est suffisant pour réguler l'expression de <i>mgtA</i>	91
Figure 27. Le peptide leader chez <i>S. enterica</i>	92
Figure 28. Importance de la séquence du peptide MgtL	94

Figure 29. Éléments potentiellement responsables du <i>calage</i>	96
Figure 30. Traduction <i>in vitro</i>	98
Figure 31. RfaH n'est pas impliqué dans le mécanisme de régulation	100
Figure 32. La RNase H n'est pas impliquée dans le mécanisme de régulation	101
Figure 33. MgtS n'est pas impliqué dans le mécanisme de régulation	102
Figure 34. Séquences requises pour la régulation	103
Figure 35. Modèle de la régulation par le riborégulateur et le peptide leader	104
Figure 36. Principe de la technique de profilage ribosomique	108
Figure 37. Distribution du nombre de lectures pour les échantillons de profilage ribosomique	110
Figure 38. Distribution du nombre de lectures en boîtes à moustaches	112
Figure 39. Répartition des ARN selon leur fonction.	112
Figure 40. Expression de <i>zinT</i>	113
Figure 41. Comparaison de la reproductibilité intra-condition	114
Figure 42. MA-plot intra-condition avant et après normalisation à la taille de l'échantillon	116
Figure 43. Ratios faible sur fort Mg^{2+} des contrôles	119

Figure 44.	Effet du Mg^{2+} sur le niveau d'ARNm de <i>corA</i>	123
Figure 45.	Résultats du profilage ribosomique pour <i>corA</i>	125
Figure 46.	Effet de la traduction de <i>corL</i> sur l'expression de <i>corA</i>	126
Figure 47.	Comparaison des transporteurs MgtA et YbeX	127
Figure 48.	Contexte génomique du gène <i>ybeX</i>	128
Figure 49.	Effet du Mg^{2+} sur le niveau d'ARNm de <i>ybeX</i>	129
Figure 50.	Effet du Mg^{2+} sur l'expression de <i>ynfB</i>	130
Figure 51.	Résultats du profilage ribosomique pour l'opéron <i>hisLGDC</i>	132
Figure 52.	Résultats du profilage ribosomique pour le gène <i>ybjX</i>	133
Figure 53.	Résultats du profilage ribosomique pour l'opéron <i>gsiAB</i>	135
Figure 54.	Résultats du profilage ribosomique pour l'opéron <i>malFEDCB</i>	137
Figure 55.	Résultats du profilage ribosomique pour l'opéron <i>cysDNC</i>	138
Figure 56.	Effet du Mg^{2+} sur l'expression de <i>ompX</i>	139
Figure 57.	Modèle de l'« embouteillage » transcriptionnel	150

CHAPITRE 1

INTRODUCTION

1.1. MISE EN CONTEXTE

1.1.1. IMPORTANCE DE L'ION Mg^{2+}

L'ion Mg^{2+} est le cation divalent le plus abondant dans les systèmes biologiques, avec une concentration cellulaire totale estimée à environ 100 mM (Moncany et Kellenberger, 1981) et une concentration libre de 1-4 mM (Froschauer et al., 2004; Lusk et al., 1968). Cette grande différence de concentration entre le Mg^{2+} lié et non lié illustre l'utilisation généralisée de cet ion responsable de la stabilité et le fonctionnement de nombreuses molécules et complexes cellulaires :

- Le Mg^{2+} est le cofacteur de nombreuses enzymes. Notamment, il est nécessaire à la liaison de l'ATP aux enzymes qui en requiert. D'après la base de données sur *Escherichia coli* ecocyc.org, plus de 300 enzymes utilisent le Mg^{2+} comme cofacteurs. Ces enzymes sont impliquées dans une large variété de fonctions cellulaires telles que la réplication de l'ADN par la polymérase d'ADN I, la maturation de certains ARN par la RNase III, la production d'énergie via le cycle de Krebs par la 2-oxoglutarate décarboxylase ou l'intégrité membranaire par la lipopolysaccharide core heptose (I) kinase.
- Il est également nécessaire au repliement des molécules d'ARN qui adoptent des structures complexes leur conférant leur fonction et leur stabilité. Alors que le repliement de la structure secondaire est facilité par la présence d'ions monovalents tels que l'ion potassium K^+ , la structure tertiaire nécessite l'ion divalent magnésium

Mg^{2+} (Pyle, 2002). Le bon repliement de la structure tertiaire est notamment important pour la fonction des ARNt (Misra et Draper, 2002). Bien que la structure devienne plus stable avec l'augmentation de Na^+ ou d'autres ions du groupe I, elle le devient dramatiquement plus en présence de Mg^{2+} . De plus une diminution du Mg^{2+} entraîne le dépliement du ribosome chez *E. coli* sans perte des protéines ribosomiques ou d'ARNr (Gesteland, 1966). La quantité de Mg^{2+} contenue dans les ribosomes représente 12 mM, soit environ 10 % du Mg^{2+} total contenu dans une cellule (Schuwirth et al., 2005). Une carence en Mg^{2+} a par ailleurs été montrée pour induire la dégradation du ribosome (McCarthy, 1962). Le Mg^{2+} est également requis pour le repliement d'ARN catalytiques appelés ribozymes. Un exemple de ribozyme est l'intron de groupe I, un ARN capable de s'exciser lui-même de son ARN précurseur. Ces ribozymes nécessitent des ions Mg^{2+} pour leur structure ainsi que pour leur activité catalytique (Rangan et Woodson, 2003).

- Le Mg^{2+} participe au maintien de l'intégrité des membranes à la fois comme cofacteur d'enzymes de biosynthèse et en formant des liens entre les charges des lipides et des protéines de la membrane (Hancock, 1984).
- Son rôle dans la virulence a été étudié, notamment, chez *Salmonella enterica* serovar Typhimurium (noté *S. enterica* dans le reste du texte). Ses gènes de virulence sont en effet sous le contrôle du système à deux composants PhoQ/PhoP qui est activé en présence de faibles concentrations de Mg^{2+} extracellulaire. Il a été proposé que la concentration intracellulaire des cellules hôtes soit basse, activant ainsi les gènes de virulence (Groisman, 1998). Le système PhoQ/PhoP sera décrit plus tard dans le paragraphe 2.6.

Pour toutes ces raisons, le Mg^{2+} est essentiel au maintien des cellules et à la croissance bactérienne (Webb, 1966). Les bactéries peuvent croître à des concentrations en Mg^{2+} plus faibles que celle du contenu cellulaire ce qui suggère la présence d'un transport actif.

1.1.2. TRANSPORTEURS DE Mg^{2+}

Plusieurs transporteurs ont été identifiés chez les procaryotes. Ceux-ci ne sont pas tous présents chez une même bactérie et présentent des structures et mécanismes très variés. Le transporteur CorA est conservé chez les bactéries et est considéré comme le système d'import de Mg^{2+} basal. Cette protéine fonctionne comme un canal multi-ionique capable de transporter sélectivement le Mg^{2+} et est également capable de transporter l'ion cobalt Co^{2+} et l'ion nickel Ni^{2+} (Maguire, 2006). Les transporteurs MgtA et MgtB sont des ATPases de type P qui permettent l'influx, mais pas l'efflux de Mg^{2+} et de Ni^{2+} . Le transfert d'un groupement phosphate de l'ATP à la protéine mène à un changement de conformation qui permet le transport du Mg^{2+} (Møller et al., 1996). Le transporteur MgtE est un pore ionique maintenu fermé par la liaison d'ion Mg^{2+} au domaine intracellulaire (Payandeh et al., 2013).

Récemment, un nouveau type de transporteur de Mg^{2+} a été identifié. Il s'agit d'un transporteur semblable aux transporteurs de la famille Nramp (pour *natural resistance-associated macrophage proteins*). Ces transporteurs d'ions métalliques sont essentiels pour de nombreux génomes bactériens comme transporteurs d'ions divalents fondamentaux (Shin et al., 2014). Ces transporteurs de Mg^{2+} putatifs ont été identifiés chez *Clostridium acetobutylicum*, encodés par les gènes *Ca_c0685* et *Ca_c3329*. Ce dernier type a été nommé NrmT (pour *Nramp-related magnesium transporter*).

Un système d'efflux a également été identifié, impliquant les protéines CorB, CorC et CorD (Gibson et al., 1991). La mutation d'un de ces gènes est associée à une augmentation de la demi-vie de l'efflux de Mg^{2+} et la mutation des trois à sa complète abolition. Bien que ce système soit très peu compris, l'identification de deux domaines dans les protéines CorB et CorC suggèrent que celles-ci sont des transporteurs permettant l'efflux du Mg^{2+} . En effet, elles contiennent un domaine CBS (ou de Bateman), également présent chez MgtE et impliqué dans

l'homéostasie du Mg^{2+} , ainsi qu'un domaine HlyC retrouvé dans de nombreuses protéines impliquées dans le transport d'ions (Zhang et al., 2010).

Chez *E. coli*, seuls les transporteurs CorA et MgtA ont été identifiés. Le transporteur CorA étant considéré comme le transporteur basal et le transporteur MgtA nécessaire à la croissance dans un milieu faible en Mg^{2+} . Par ailleurs, le gène *ybeX* présente des similitudes avec *corC*, et pourrait donc coder pour un système d'efflux du Mg^{2+} .

L'importance du Mg^{2+} et la présence nécessaire de transporteurs spécifiques suggèrent que son homéostasie chez les bactéries est essentielle. Le maintien d'une concentration intracellulaire stable, quelle que soit la concentration extracellulaire requiert ainsi la régulation de la quantité et l'activité des transporteurs de Mg^{2+} .

1.2. RÉGULATION DE L'EXPRESSION GÉNIQUE

1.2.1. EXPRESSION GÉNIQUE

L'expression génique est le processus complexe par lequel l'information contenue dans l'ADN est décodée en protéines. Le niveau d'expression d'un gène dépend de la transcription, la traduction et la dégradation de l'ARNm.

1.2.1.1. Transcription

Dans un premier temps, l'ADN est transcrit en ARNm lors de la transcription. Ce processus consiste à générer un brin d'ARN complémentaire à la matrice d'ADN. Cette réaction est

catalysée par la polymérase d'ARN (PdARN). L'apoenzyme PdARN chez *E. coli* est composée de 5 sous unités ($\alpha_2\beta'\beta\omega$) (Figure 1). La transcription se fait en trois étapes : l'initiation, l'élongation et la terminaison.

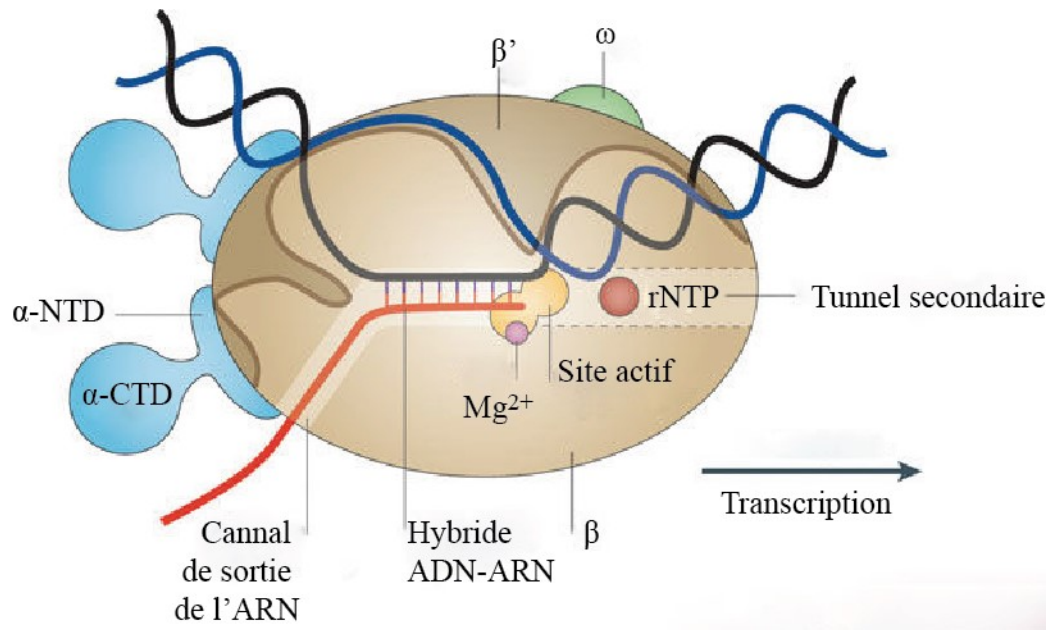


Figure 1. Le cœur de la polymérase d'ARN avec ses sous unités α , β et β' et ω . Le brin matrice est représenté en noir, le brin non matrice en bleu. L'ARN naissant est en rouge. Les domaines terminaux α -aminés (α -NTD) permettent l'assemblage du complexe; les domaines terminaux α -carboxylés (α -CTD) et la sous unité ω ont un rôle régulateur lors de l'initiation. Les sous unités β et β' forment ensemble le site actif et font les contacts avec les acides nucléiques. Le ribonucléotide 5'-triphosphate (rNTP) lié à un ion Mg^{2+} entrerait par le tunnel secondaire. L'ARN transcrit est évacué dans le tunnel de sortie. Adapté de Santangelo et Artsimovitch, 2011.

L'initiation de la transcription ne peut avoir lieu qu'à des régions de l'ADN précises que sont les promoteurs. Un promoteur est composé de deux éléments de séquence très conservée, les boîtes -10 et -35 (positions par rapport au premier nucléotide transcrit). Pour l'initiation la PdARN requiert la liaison d'un facteur σ qui permet à cette holoenzyme de reconnaître les régions promotrices. Il existe différents facteurs σ qui sont spécifiques à différents promoteurs (*cf.* paragraphe 2.2.1). La liaison spécifique au promoteur forme un complexe initial fermé et induit une série d'importants changements conformationnels qui plie le duplex d'ADN en aval

dans le site actif de la PdARN puis ouvre par la suite l'ADN pour former la bulle de transcription, plaçant le brin matrice de l'ADN dans le site actif (Ruff et al., 2015).

Une fois la bulle de transcription formée, le facteur σ se détache de la PdARN qui peut alors commencer l'élongation. Cette réaction de polymérisation se fait par la formation d'un pont phosphodiester qui s'établit entre le nucléotide à ajouter dans le site actif de la PdARN et l'hydroxyde en 3' de l'ARN. L'ARN est par la suite transloqué pour permettre l'ajout du nucléotide suivant.

Bien que le complexe d'élongation soit très stable et puisse transcrire de grandes régions d'ADN (jusqu'à 10^5 paires de bases chez les bactéries (Santangelo et Artsimovitch, 2011)), la progression de la PdARN est influencée par des signaux variés tels que les pauses de la PdARN ou des facteurs de transcription. Une pause transcriptionnelle est un arrêt temporaire de la PdARN sur l'ADN. Elle se caractérise par un changement de conformation de la PdARN dans un état de pause élémentaire qui est à l'origine de différents types de pause et peut conduire à l'arrêt, la terminaison ou la reprise de l'élongation de la transcription (Weixlbaumer et al., 2013).

Un arrêt de la transcription peut se produire lorsque la PdARN s'interrompt dans une conformation qui ne lui permet pas de reprendre spontanément l'élongation en absence du clivage de l'ARN. Cet arrêt fait suite au recul de la PdARN sur l'ADN et l'ARN (Komissarova et Kashlev, 1997). Pour sortir de l'état d'arrêt, la PdARN doit cliver l'ARN grâce à son activité nucléolytique intrinsèque induite par GreA ou GreB. Après la dissociation du fragment d'ARN 3', la transcription peut reprendre à partir du nouveau 3' du transcrit (Surratt et al., 1991).

À la fin d'un transcrit, des régions terminatrices signalent à la PdARN qu'elle doit relâcher l'ARN en se dissociant de l'ADN et l'ARN. Il existe deux types de terminateurs : les terminateurs intrinsèques et les terminateurs dépendant de Rho.

Aux sites intrinsèques, c'est un signal d'acides nucléiques qui induit la dissociation du complexe d'élongation. La reconnaissance de ces signaux par le complexe d'élongation ne nécessite aucun facteur, mais peut être améliorée par des protéines accessoires, tel que le facteur général d'élongation de la transcription NusA (Schmidt et Chamberlin, 1987). Un terminateur intrinsèque canonique a été décrit en détail chez *E. coli*. C'est un signal d'ARN composé d'une tige-boucle d'ARN riche en GC suivie par une queue poly-U. La terminaison s'effectue en deux étapes : la pause de la PdARN au niveau de la queue poly-U puis la dissociation de l'ARN (Gusarov et Nudler, 1999). Des terminateurs intrinsèques non canoniques ont également été identifiés, mais l'absence de caractéristiques communes rend leur prédiction très difficile (Mitra et al., 2009).

La terminaison aux terminateurs dépendant de facteurs requiert l'action d'une protéine régulatrice telle que Rho (Roberts, 1969). Rho est une protéine hexamérique qui possède une activité ATPase dépendante de l'ARN (Galluppi et Richardson, 1980). Rho lie initialement l'ARN au niveau d'une séquence *rut* (*rho utilization*). Ces séquences d'environ 80 nucléotides sont connues pour être peu conservées, riches en cytosines et peu structurées (Ciampi, 2006). Rho interagit avec le site *rut* au niveau de ses sites de contact primaire sur chaque sous-unité, ce qui mène l'ARN à être enroulé autour de l'hexamère de Rho. Rho se referme alors sur l'ARN et fait des contacts additionnels grâce au site secondaire au niveau du centre de l'hexamère. Rho se transloque par la suite le long de l'ARN de 5' en 3'. Plusieurs modèles de translocation ont été proposés, mais aucun consensus n'a été atteint. Cependant, il est admis que Rho utilise l'hydrolyse de l'ATP pour ses activités de translocation et d'hélicase qui permettent le déroulement du complexe ADN:ARN pour libérer l'ARN. La capacité de

translocation de Rho lui permet d'éventuellement rattraper la PdARN pour terminer la transcription. Le plus fréquemment les sites de terminaison de Rho sont des sites où la PdARN pause. Cependant, tous les sites de pauses ne sont pas des sites de terminaisons (Banerjee et al., 2006). Rho est essentiel chez *E. coli* et est responsable de la terminaison de 70% des transcrits, aussi bien des ARNm, des ARN de transfert (ARNt) et des petits ARN (pARN) (Peters et al., 2012). Il est notamment important pour empêcher la transcription antisens, prévenir la transcription de tous les gènes d'un opéron et aide à enlever les R-loops (Harinarayanan et Gowrishankar, 2003).

Le complexe d'élongation ne se compose pas seulement de la PdARN, de l'ARN et de la matrice d'ADN. Des protéines accessoires ou facteurs de transcriptions tels que les facteurs GreA/GreB ou Nus peuvent s'y associer et moduler l'élongation de la transcription (*cf.* paragraphe 2.2.1).

1.2.1.2. Traduction

Chez les procaryotes, la transcription est couplée à la traduction. Pour la majorité des gènes bactériens, la traduction est initiée très rapidement après que le site de reconnaissance du ribosome émerge du canal de sortie de la PdARN. Le premier ribosome suivant la PdARN assiste ainsi celle-ci directement pendant l'élongation de la transcription. Ceci se traduit par des vitesses de transcription et de traduction couplées, soit un acide aminé traduit pour trois nucléotides transcrits (Proshkin et al., 2010). La présence de ribosome empêche ainsi le recul (*backtracking*) de la PdARN. Une étude récente a confirmé l'existence d'un complexe structural entre la PdARN et le ribosome, appelé *expressome* (Kohler et al., 2017). Le couplage de la transcription et de la traduction permet notamment d'empêcher la formation de R-loops (Gowrishankar et Harinarayanan, 2004). Une *R-loop* est une structure d'acides nucléiques à trois brins, composée d'un hybride ADN:ARN et du brin d'ADN non matrice

associé. La formation de *R-loops* est associée à une instabilité génomique notamment due à une mutagenèse spontanée importante. Chez *E. coli*, les *R-loops* sont défaits par la RNase H qui dégrade les ARN impliqués dans des hybrides ADN:ARN (Aguilera et García-Muse, 2012).

La traduction est l'étape qui permet la synthèse protéique à partir de l'ARNm, C'est une réaction de polymérisation d'acides aminés. La machinerie de traduction est appelée ribosome. Celui-ci contient trois ARNr, 5S, 16S, 23S, ainsi que des protéines ribosomiques. Il est composé de deux sous-unités appelées 30S pour la petite et 50S pour la grande, selon leur coefficient de sédimentation. Le 30S est composé de l'ARNr 16S et de 21 protéines ribosomiques (S1 à S21) (Culver et Kirthi, 2008). Le 50S est composé des ARNr 23S et 5S et de 33 protéines ribosomiques (Shajani et al., 2011). Lorsque les deux sous-unités sont assemblées, le complexe alors est appelé 70S. Différents sites fonctionnels ont été identifiés dans le complexe traductionnel formé par le ribosome et l'ARNm (Figure 2). Le site A qui lie les ARNt-aminoacylés (ARNt-aa : ARNt liés à un acide aminé) et le site P qui lie les ARNt-peptidyl (ARNt liés au peptide en cours de synthèse) se situent au niveau de la grande sous-unité, contenus au niveau du centre peptidyl-transférase (peptidyl-transferase center (PTC)) qui catalyse la formation du lien peptidique. Il y a également le site E où le ribosome lie les ARNt déacylés avant qu'ils quittent le ribosome. Enfin il y a le tunnel de sortie par lequel le polypeptide naissant sort du ribosome (Steitz, 2008).

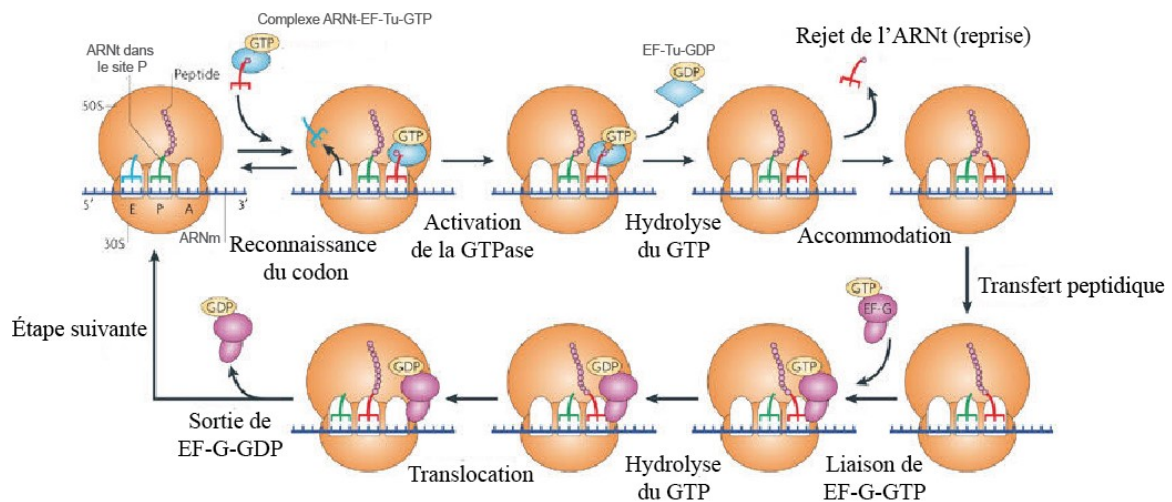


Figure 2. Élongation de la traduction. Un ARNt est apporté au site A en complexe avec le facteur d'élongation EF-Tu lié à une molécule de GTP (EF-Tu-GTP). L'appariement correct codon-anticodon active le centre GTPase du ribosome qui permet l'hydrolyse de la GTP et la libération de l'extrémité aminoacyl de l'ARNt de EF-Tu. La liaison de l'ARNt induit également un changement conformationnel de l'ARNr qui oriente de manière optimale l'ARNt-peptidyl et l'ARNt-aminoacyl pour que la réaction peptidyl-transférase ait lieu. La chaîne peptidique est alors transférée sur l'ARNt au site A. Le ribosome doit par la suite se déplacer dans la direction 5' → 3' de l'ARNm de telle sorte qu'il puisse décoder le codon suivant. La translocation des ARNt et de l'ARNm est facilitée par la liaison de la GTPase EF-G qui permet à l'ARNt déacylé au site P de se déplacer vers le site E et à l'ARNt peptidyl au site A de se déplacer au site P grâce à l'hydrolyse du GTP. Le ribosome est ensuite prêt pour une autre étape d'élongation de la traduction. L'ARNt déacylé dans le site E est alors relâché lors de la liaison de l'ARN aminoacyl suivant au site A. L'élongation se termine lorsqu'un codon stop est atteint, ce qui initie la terminaison de la traduction et la libération du polypeptide. Adapté de Steitz, 2008.

La phase d'initiation de la traduction commence par la formation du complexe d'initiation 30S (CI30S) dans lequel le codon d'initiation de l'ARNm (AUG, GUG ou UUG) est décodé par l'anticodon CAU de l'ARNt initiateur fMet (formylméthionine) dans le site P de la sous-unité 30S du ribosome. Le CI30S est ensuite rejoint par la sous-unité 50S pour former le complexe d'initiation 70S (CI70S), capable de former un dipeptide initiateur avec l'ARNt-aa encodé par le deuxième codon de l'ARNm apporté au site A du ribosome par le facteur d'élongation EF-Tu. Trois facteurs (les facteurs d'initiation IF1, IF3 et IF3) assistent la formation du CI30S et s'en dissocient lors de la transition du CI30S au CI70S. Une caractéristique importante de la reconnaissance de la région d'initiation de la traduction des ARNm bactériens est la présence d'une séquence Shine-Dalgarno (SD) complémentaire à la région 3' de l'ARNr 16S (anti-SD)

située en amont du codon d'initiation, idéalement entre 4 et 9 nucléotides. Cependant le SD n'est ni suffisant ni nécessaire à l'initiation de la traduction : des ARNm sans SD peuvent être traduits et la présence d'un SD en amont d'un codon d'initiation n'est pas toujours associée à un événement d'initiation de la traduction. (Gualerzi et Pon, 2015).

L'élongation de la traduction se passe en trois étapes : (1) la liaison d'un tRNA-aa à son codon correspondant présent au site A du ribosome (2) la formation du lien peptidique entre l'ARNt-aa du site A et l'ARNt-peptidyl dans le site P, et (3) la translocation des deux ARN du site A vers le site P et du site P vers le site E. Les réactions de décodage et de translocation sont assistées par les facteurs d'élongation (EF) : EF-Tu accélère la liaison de l'ARNt-aa au ribosome, EF-Ts est le facteur d'échange de nucléotide guanine pour EF-Tu, et EF-G favorise la translocation des ARNt.

L'élongation de la traduction n'est pas cinétiquement constante, avec des périodes de synthèse rapide séparées par des pauses (Li et al., 2012). Les raisons pour lesquelles la traduction pause sont diverses et mal comprises. Une pause peut être causée par exemple par des codons rares, des séquences SD internes au cadre de lecture, ou par des structures secondaires de l'ARNm. Le peptide naissant du canal de sortie (qui peut accommoder entre 30 et 40 acides aminés (Voss et al., 2006)) peut également interagir avec le ribosome et contrôler ainsi la traduction (Wilson et al., 2016). Bien qu'aucune définition exacte n'existe dans la littérature, le *calage* du ribosome décrira ici une pause longue avec un potentiel régulateur. Une pause qui ne permet pas au ribosome de reprendre la traduction est nommée arrêt et diffère de la terminaison de la traduction par le mécanisme qui permet de recycler les ribosomes.

La terminaison de la traduction a lieu lorsqu'un codon de terminaison ou stop (UGA, UAG or UAA chez *E. coli*) est rencontré au site A du ribosome. La libération de la chaîne polypeptidique naissante du tRNA-peptidyl se situe au site P du ribosome lorsque le site A est

occupé par un des facteurs de terminaison (RF1, RF2 ou RF3) qui interagissent avec les codons stop. Les facteurs RF de classe I, soit RF1 et RF2, remplacent les ARNt pour la reconnaissance du codon et déclenchent l'hydrolyse du tRNA-peptidyl. Le facteur de classe II RF3 est une protéine liant le GTP qui stimule la dissociation de RF1 ou RF2 du ribosome après le détachement du peptide, étape nécessitant l'hydrolyse du GTP (Freistroffer et al., 1997). Le complexe de terminaison est par la suite désassemblé grâce à deux facteurs: le facteur de recyclage du ribosome (RRF) et le facteur d'élongation EF-G (Janosi et al., 1996).

1.2.1.3. Dégradation de l'ARNm

Un dernier facteur déterminant du niveau d'expression d'un gène est la dégradation de l'ARNm. Une dégradation rapide permet d'ajuster continuellement la population de messagers aux besoins de la bactérie en protéines spécifiques. Les ARNm ont ainsi une demi-vie courte, qui peut aller de quelques secondes à plus d'une heure, mais d'environ 7 minutes en moyenne (Selinger et al., 2003). Le processus complet de dégradation d'un ARN pleine longueur en acides nucléiques libres peut être décomposé en deux étapes : l'initiation et la dégradation des intermédiaires produits (Deutscher, 2006) (Figure 3). De manière générale la dégradation est initiée par une endoribonucléase qui clive l'ARN à des sites spécifiques internes. Par la suite, ce sont des exoribonucléases qui vont dégrader les fragments d'ARN produits par l'endoribonucléase dans le sens 3' → 5'. Les endoribonucléases clivent l'ARN par hydrolyse d'un lien phosphodiester produisant ainsi deux fragments d'ARN. Elles ne reconnaissent pas nécessairement une séquence spécifique, mais peuvent également reconnaître des structures d'ARN.

La RNase E est considérée comme l'enzyme responsable de la majorité de la dégradation chez *E. coli* (Stead et al., 2011). Elle est spécifique aux régions simple-brin avec une préférence pour les séquences riches en A et U (McDowall et al., 1994). Il existe deux voies d'entrée pour

la RNase E dépendamment du nombre de phosphates en 5'. La présence d'un 5' monophosphate stimule l'activité de la RNase dans la région 5'. La conversion du groupe triphosphate en monophosphate est réalisée par la protéine RppH (Deana et al., 2008). Cependant certains substrats structurés peuvent être clivés avec leur 5' triphosphate intact, ce qui constitue le deuxième mécanisme appelé entrée directe (Kime et al., 2009).

La RNase G est un paralogue de la RNase E (McDowall et al., 1993) et est également impliquée dans la dégradation des ARN (Carpousis et al., 2009). Ces deux enzymes partagent une haute identité (35 %) et similarité (50 %) de séquence et reconnaissent des substrats similaires riches en AU, mais leurs sites de clivage précis ne semblent pas parfaitement conservés (Tock et al., 2000).

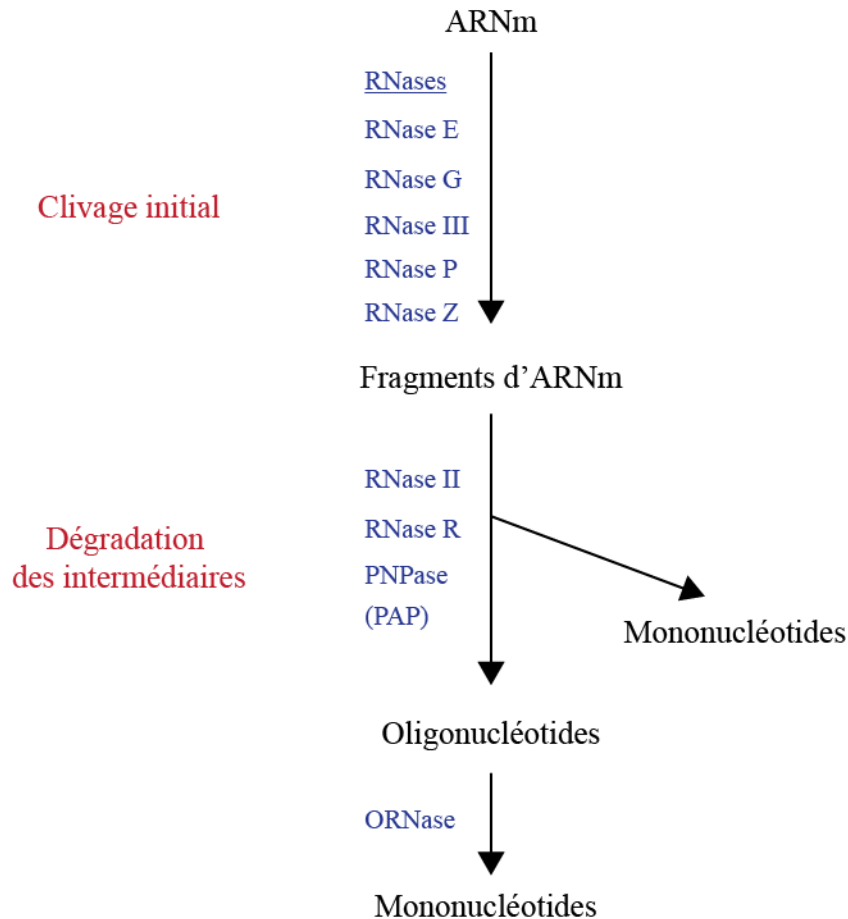


Figure 3. La dégradation des ARNm chez *E. coli*. Elle peut être induite par plusieurs endoribonucléases qui produisent des fragments intermédiaires qui vont être pris en charge par des exoribonucléases. La polymérase poly(A) (PAP) peut ajouter une queue poly(A) en 3' afin de faciliter la dégradation par ces enzymes. Enfin les petits fragments d'ARN produits sont dégradés en mononucléotides par l'oligoribonucléase (ORNase). PNPase : phosphorylase polynucléotide. Inspiré de Deutscher, 2006.

La RNase III est hydrolytique et est spécifique aux ARN double-brin générant un 5'-monophosphate et un 3'-hydroxyl avec un bout flottant de deux bases (Meng et Nicholson, 2008). Elle est responsable de la dégradation de plusieurs ARNm (Calin-Jageman et Nicholson, 2003; Lim et al., 2012), en particulier elle initie la dégradation de l'ARNm du gène *pnp* encodant la phosphorylase polynucléotide (PNPase) (Régnier et Portier, 1986).

La RNase P est une ribonucléoprotéine. Ce ribozyme est responsable de la maturation des ARNt et est connu pour cliver certains ARNm tels que l'ARNm *btuB* chez *E. coli* (Altman et al., 2005) et pourrait ainsi induire leur dégradation.

La RNase Z chez *E. coli* est capable de cliver plusieurs ARNm (Perwez et Kushner, 2006) et pourrait donc induire leur dégradation.

Après clivage initial par une de ces endoribonucléases, l'ARNm est par la suite dégradé par des exoribonucléases : la PNPase, la RNase II et/ou la RNase R.

La PNPase catalyse processivement la dégradation de l'ARN dans le sens 3'→5'. Bien que l'activité de la PNPase chez *E. coli* soit connue pour être bloquée par les structures d'ARN double brin (Spickler et Mackie, 2000), cette RNase peut s'associer à d'autres enzymes telles que l'hélicase d'ARN B (RhlB) pour dégrader les ARN structurés (Lin et al., 2005). Une région 3' non appariée de 7 à 10 nucléotides minimum est requise pour que la PNPase puisse s'y lier (Cheng et Deutscher, 2005).

La RNase II de *E. coli* est une exoribonucléase hydrolytique qui dégrade l'ARN dans le sens 3'→5' de manière indépendante de la séquence, générant des nucléosides 5'-monophosphate. Cette enzyme est connue pour s'arrêter environ 7 nucléotides en amont d'une région d'ARN double brin (Spickler et Mackie, 2000).

La RNase R est une exoribonucléase hydrolytique qui dégrade dans le sens 3'→5'. Elle est unique parmi les exonucléases de *E. coli* grâce à sa capacité de dégrader facilement les ARN hautement structurés (Awano et al., 2010).

Contrairement à la RNase II, la PNPase et la RNase R peuvent donc dégrader des ARN structurés, mais nécessitent néanmoins une région non appariée pour initier la dégradation. Afin d'aider ces RNases à dégrader l'ARN, une enzyme ajoute une vingtaine d'adénines en 3' (ou queue poly(A)) : la polymérase poly(A) ou PAP (Belasco, 2010). La PAP permet notamment la dégradation des tiges-boucle que sont les terminateurs de transcription par les exoribonucléases.

La combinaison des enzymes décrites précédemment mène à l'accumulation de fragments d'ARN de 2-5 nucléotides qui peuvent être délétères pour la cellule (Ghosh et Deutscher, 1999). L'oligoribonucléase est l'enzyme qui dégrade ces oligoribonucléotides courts en nucléotides simples prêts à être réutilisés pour la transcription de nouveaux ARN (Niyogi et Datta, 1975).

Le dégradosome est un large complexe multi-protéique impliqué dans la dégradation des ARN (Bernstein et al., 2004). Il est convenu que celui-ci agit en tant que machine de dégradation générale des ARN, dans laquelle les composants coopèrent. La formation du complexe contribue à la coordination du clivage endoribonucléolytique avec la dégradation exoribonucléolytique. Chez *E. coli*, ce complexe se compose de l'endoribonucléase RNase E, de l'exoribonucléase PNPase, de l'hélicase RhlB et de l'enzyme glycolytique énalase (Vanzo et al., 1998). Le complexe s'assemble sur la région C-terminal de la RNase E. D'autres composants mineurs peuvent s'associer au dégradosome et moduler son activité enzymatique, telles que la kinase polyphosphate, la PAP, des protéines ribosomiques, les chaperonnes DNaK et GroEL et d'autres hélicases (Regonesi et al., 2006). La composition du dégradosome peut également varier en fonction des conditions de croissance ou en cas de stress. Par exemple, un complexe différent, contenant la RNase E, Hfq et le petit ARN régulateur SgrS, est formé en cas de stress sucre-phosphate (Morita et al., 2005).

Les bactéries vivent dans des environnements sujets à des variations rapides, la plupart étant sous-optimaux pour la croissance, et utilisent une variété de mécanismes de transduction du signal pour répondre rapidement à ces conditions changeantes. Pour cela, l'expression génétique est contrôlée et régulée finement à plusieurs niveaux : au niveau de l'initiation et de l'élongation de la transcription, de la stabilité de l'acide ribonucléique messenger (ARNm), de l'initiation et de l'élongation de la traduction et de la stabilité et de modifications post-traductionnelles de la protéine.

1.2.2. LA RÉGULATION DE L'EXPRESSION GÉNÉTIQUE

1.2.2.1 Transcription

Pour la plupart des gènes bactériens, l'étape clef de la régulation semble la reconnaissance du promoteur et de l'initiation de la transcription par la PdARN. L'initiation de la transcription peut être régulée au niveau de la formation de l'holoenzyme, de la reconnaissance du promoteur ou de l'activité de la polymérase. Ces mécanismes peuvent soit moduler la fonction de la polymérase soit moduler l'accessibilité ou l'affinité pour la région promotrice. Bien que le facteur σ principal (σ^{70}) recrute la polymérase à la majorité des promoteurs, *E. coli* possède plusieurs autres facteurs σ dits alternatifs pour guider la polymérase à différents promoteurs. Ces facteurs sont exprimés dans des conditions spécifiques (stress, phase de croissance, etc.) et permettent la transcription d'un nombre limité de gènes (Gruber et Gross, 2003). D'autres facteurs sont également capables d'interagir avec la polymérase afin de modifier son activité comme la protéine DksA qui peut moduler la stabilité du complexe d'initiation en cas de stress métabolique grâce au ppGpp par exemple (Ross et al., 2013). D'autres facteurs peuvent diminuer le nombre de polymérases disponibles pour la transcription en séquestrant celle-ci. Notamment l'ARN non codant 6S, transcrit en réponse à une croissance lente, imite les promoteurs cibles de la polymérase liée au facteur σ^{70} et forme un complexe avec celle-ci pour

l'empêcher de se lier à l'ADN. Il en résulte une augmentation de la proportion de polymérase liée au facteur σ^{38} , reprogrammant ainsi le transcriptome (Cavanagh et Wassarman, 2014).

La régulation de l'initiation de la transcription peut également s'effectuer par la liaison de facteurs directement au promoteur (Dorman, 2013; Lee et al., 2012; Ruff et al., 2015). Une telle régulation peut faire intervenir divers mécanismes tels que le surenroulement de l'ADN ou le blocage de l'accès du promoteur à la polymérase. Dans ce cas, les facteurs de transcription ont des domaines capables de reconnaître spécifiquement des motifs d'ADN. Bien que certains régulateurs ne ciblent qu'un promoteur, beaucoup régulent de nombreux promoteurs et à l'inverse beaucoup de promoteurs sont régulés par plusieurs facteurs (Martínez-Antonio et al., 2008). La régulation au promoteur génère ainsi un réseau de régulation complexe qui permet une régulation spécifique à chaque type d'environnement. Les promoteurs peuvent également être modifiés eux-mêmes, le mécanisme le plus courant étant la méthylation de l'ADN, modifiant ainsi l'accessibilité à la PdARN ou aux facteurs de transcription (Casadesus et Low, 2006).

Bien que réguler l'initiation de la transcription soit suffisant pour réguler l'expression génique, toutes les autres étapes suivantes peuvent également être régulées.

L'élongation de la transcription dépend de signaux nucléiques (ADN et/ou ARN) et de la présence de protéines qui se lient à l'ADN, à l'ARN ou à la PdARN. L'atténuation de la transcription correspond à une terminaison prématurée de la transcription entraînant une diminution de l'expression. Celle-ci peut être régulée par la présence de ribosome, de protéines ou d'ARNt *via* la présence d'un peptide leader (voir paragraphe sur les peptides leaders), ou par des petites molécules *via* la présence d'un riborégulateur (voir paragraphe sur les riborégulateurs).

1.2.2.2. Dégradation de l'ARNm

La régulation de la dégradation des ARN peut être indirecte suite à l'inhibition de la traduction ou directe indépendamment de l'état de traduction (Deana et Belasco, 2005). Les signaux déterminant si un ARNm doit être dégradé sont variés : la présence d'une structure (Lim et al., 2012), la modulation de l'accès au site de clivage par un riborégulateur (Caron et al., 2012) (voir paragraphe 2.4 sur les riborégulateurs) ou la liaison d'un petit ARN (pARN (Caron et al., 2010) (voir paragraphe 2.3 sur les pARN). Certains processus de maturation peuvent à l'inverse stabiliser l'ARNm (Rochat et al., 2013).

1.2.2.3. Traduction

Bien que la régulation du niveau d'ARNm contrôle indirectement la quantité de protéines, la concentration d'ARNm seule n'explique qu'une part de la variation du nombre de protéines (environ la moitié chez *E. coli* (Corbin et al., 2003; Lu et al., 2007b)). La régulation de la traduction est par ailleurs critique puisque la bactérie en croissance rapide utilise jusqu'à 40 % de son énergie totale pour la synthèse protéique chez *E. coli* (Gold, 1988). La traduction peut être régulée au niveau de l'initiation, l'élongation et la terminaison.

L'initiation de la traduction peut être contrôlée par des éléments de l'ARNm, des protéines ou des ARN exprimés en *trans*. Bien que le ribosome possède une activité hélicase pour défaire les structures d'ARNm (Takyar et al., 2005), des structures au niveau du SD sont connues pour limiter l'accessibilité au ribosome (de Smit et van Duin, 1990). La modulation de ces structures en présence d'un ligand ou de la température par exemple permet de réguler la traduction en fonction des conditions (voir paragraphe 2.4 sur les riborégulateurs). En plus du 5'UTR, plusieurs exemples montrent que la région 3'UTR peut également avoir un impact sur l'initiation de la traduction soit en s'appariant directement au 5'UTR (Balaban et Novick,

1995; Ruiz de los Mozos et al., 2013) soit par réarrangement structural suite à une étape de maturation (Franch et al., 1997). Le SD n'est pas la seule cible de la régulation de l'initiation traductionnelle. En effet, plusieurs études ont mis en évidence la présence de sites « d'attente » ou standby du ribosome en amont du SD permettant la traduction d'ARNm dont le SD est séquestré. Une étude récente a découvert un mécanisme par lequel un ARN antisens s'hybride à un tel site sur l'ARNm *tisB* empêchant la traduction de ce dernier dont le SD est hautement séquestré (Darfeuille et al., 2007). Certains ARNm traduits ne contiennent étonnamment pas de séquence SD, ni même de région 5'. Ces ARNm sont appelés *leaderless* et nécessitent dans ce cas la formation du complexe 30S-fMet-tRNA_f^{Met}-IF2 (Grill et al., 2000). Mis à part les éléments d'ARN présents, le recrutement de protéines influence également l'initiation de la traduction. Par exemple, les protéines ThRS et CsrA bloquent la liaison du ribosome en se liant de façon compétitive au SD (Baker et al., 2002; Moine et al., 1990), la protéine S15 prévient la formation d'un complexe 30S-mRNA- tRNA_f^{Met} actif en se liant à la région 5' de son ARNm (Philippe et al., 1993) et à l'inverse, la protéine S1 aide le 30S à se lier et à défaire les structures au niveau de son site de liaison (Duval et al., 2013).

Bien que la majorité des mécanismes de régulation de la traduction connus affectent l'étape d'initiation, l'élongation de la traduction peut également être régulée. Bien que l'importance du biais d'usage des codons soit controversée (Guimaraes et al., 2014; Nie et al., 2006), plusieurs études ont montré une corrélation entre la régulation de l'efficacité de la traduction et l'abondance relative en ARNt, par exemple en réponse à un changement de disponibilité en oxygène O₂ (Taylor et al., 2013).

Une fois la traduction complétée, des modifications post-traductionnelles et la dégradation des protéines peuvent également affecter le niveau d'expression effectif d'un gène. Les modifications post-traductionnelles peuvent permettre le bon repliement de la protéine, la localisation cellulaire de celle-ci, ou la liaison réversible d'une molécule. Le renouvellement

ou *turnover* des protéines n'aurait cependant pas un impact important en comparaison aux autres étapes régulées (Maier et al., 2011).

Bien que pendant des décennies les protéines ont été considérées comme étant les acteurs majeurs de la régulation de l'expression génique, il est devenu évident que la régulation par des ARN régulateurs est au moins aussi importante, avec la découverte d'un nombre important de gènes régulés post-transcriptionnellement et celle de régulateurs globaux (Nogueira et Springer, 2000). Les deux types d'ARN régulateurs présents chez les bactéries sont les petits ARN (pARN) et les riborégulateurs (ou *riboswitches*) (Waters et Storz, 2009).

1.2.3. LES PETITS ARN

Les pARN sont des ARN courts exprimés en *trans* qui ne sont généralement pas traduits et qui contrôlent l'expression de leurs cibles ARNm en s'y appariant. Ils sont présents dans tous les règnes du vivant et interviennent dans différents processus biologiques importants tels que le métabolisme énergétique, le *quorum sensing*, la formation de biofilm, la réponse au stress, l'adaptation aux conditions de culture et la virulence (Michaux et al., 2014). Ils ne sont généralement pas traduits et peuvent réguler l'expression de leur cible en contrôlant leur traduction, leur transcription ou en modulant leur stabilité (Michaux et al., 2014). Les mécanismes de régulation sont variés (Figure 4) et dans beaucoup de cas, la liaison et la régulation par les pARN est facilitée par la liaison de la chaperonne Hfq (Vogel et Luisi, 2011). Ils peuvent contrôler la traduction en s'appariant jusqu'à 70 nucléotides en amont ou 15 nucléotides en aval du codon d'initiation (AUG). Dans certains cas, c'est le petit ARN qui va masquer le SD (Figure 4A) et dans d'autres c'est Hfq (Figure 4B). Un pARN peut même réguler la traduction loin du SD et de l'AUG en modulant la structure de l'ARN (Figure 4C). D'autres mécanismes ont été décrits récemment, tels que le cas d'un pARN qui déstabilise une structure en aval de l'AUG et inhibe la traduction (Jagodnik et al., 2017). Certains pARN

peuvent servir de leurre pour séquestrer un autre pARN (Figure 4E) ou une protéine régulatrice (Figure 4D). Pour moduler la stabilité de l'ARN, les pARN peuvent moduler l'accessibilité à un site de clivage directement ou non (Figures 4G et I) ou indirectement en contrôlant la traduction, l'absence de ribosome sur le cadre de lecture laissant l'ARN susceptible à la RNase E (Figure 4H). Ils peuvent enfin réguler la transcription en modulant la formation d'un terminateur intrinsèque de transcription ou l'accessibilité du site *rut* (Figure 4 J et K).

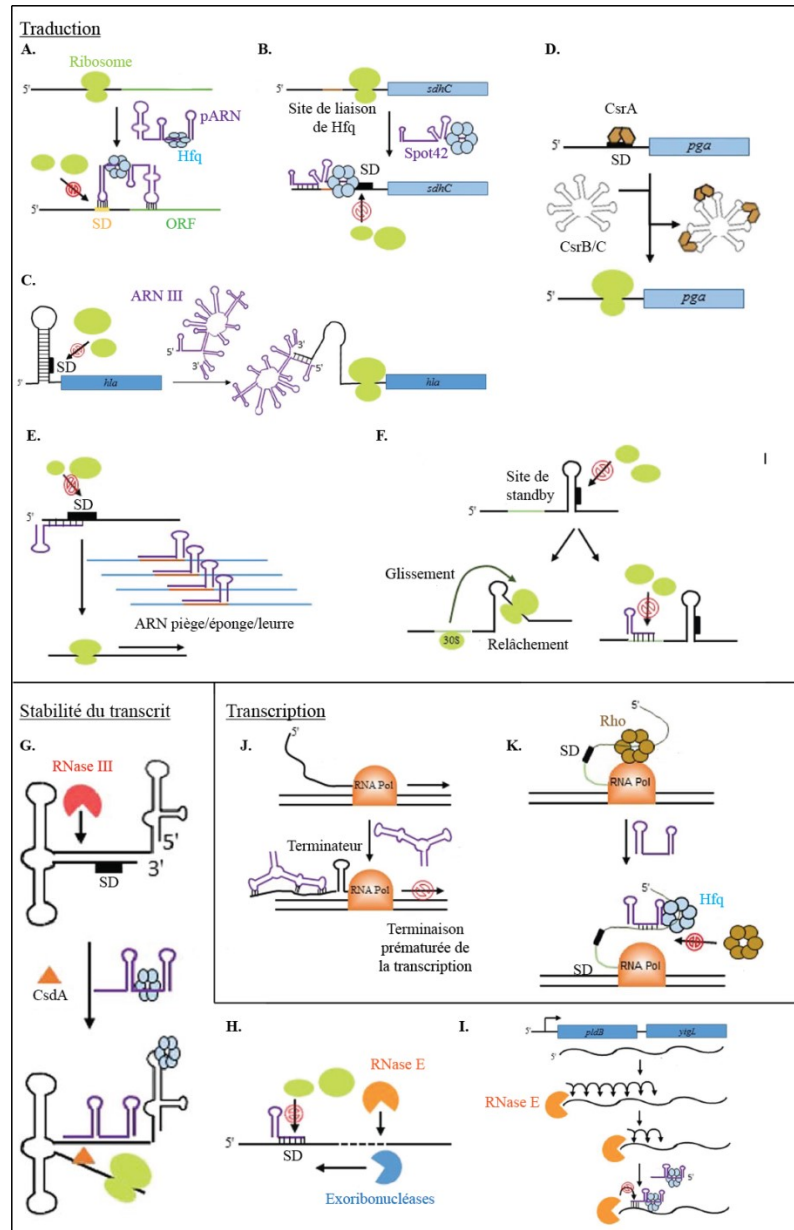


Figure 4. Mécanismes de régulation des pARN. A-F mécanismes de régulation de la traduction. A) Le pARN, lie plus qu'une région de sa cible et interfère directement avec le ribosome. B) Le pARN recrute Hfq pour bloquer le SD de sa cible. C) Le pARN ARN III chez *Staphylococcus aureus* s'apparie au 5'-UTR de l'ARNm *hla* ARNm, libérant son SD pour la traduction. D) Les pARN CsrB/C titrent la protéine CsrA et enlèvent ainsi la répression de l'expression de *pga*. E) Les pièges ARN, avec une région complémentaire à la séquence du pARN compétitionnent avec l'ARNm cible pour la liaison du pARN. F) Le pARN interfère avec la liaison du 30S au site de standby, inhibant ainsi la traduction. G) Un pARN protège le site de clivage de la RNase III et facilite la traduction avec l'aide de l'hélicase CsdA. H) L'inhibition de la traduction par un pARN est suivie par une dégradation du transcrit par la RNase E. I) L'appariement du pARN protège le transcrit en cachant un site de clivage par la RNase E. J) L'appariement du pARN induit la formation d'un terminateur prématuré de la

transcription. K) Le pARN induit interfère avec la liaison de Rho et permet la transcription complète. Adapté de Dutta et Srivastava, 2018.

1.2.4. LES RIBORÉGULATEURS

Les riborégulateurs ou « riboswitches » sont des éléments d'ARN structurés retrouvés majoritairement dans les régions 5' non traduites des ARNm bactériens et capables de réguler l'expression génique co- ou post-transcriptionnellement. Cette régulation est dépendante de la liaison spécifique d'un petit ligand non protéique qui induit un changement de structure de l'ARN (Nudler et Mironov, 2004). Les gènes régulés par des riborégulateurs sont généralement impliqués dans la biogenèse ou le transport du ligand reconnu.

Suite à leur découverte il y a une dizaine d'années (Mironov et al., 2002 ; Nahvi et al., 2002 ; Winkler et al., 2002a et b), les riborégulateurs sont désormais reconnus pour être des éléments importants et répandus de l'expression génique chez de nombreuses bactéries éloignées, et possédant des homologues chez les *archae*, les *plantae*, les *fungi* et les *algae*.

Un riborégulateur contrôle l'expression génique en changeant de conformation suite à la reconnaissance d'un signal cellulaire tel que la présence d'un métabolite cellulaire, d'un ion ou d'un stress physico-chimique (Bastet et al., 2011). Ils sont majoritairement retrouvés dans des régions 5' non traduites des ARNm bactériens et régulent en *cis* le gène situé en aval sur la même molécule d'ARNm. Ils sont composés de deux domaines fonctionnels, l'aptamère et la plateforme d'expression (Bastet et al., 2011). L'aptamère est le domaine qui permet de détecter spécifiquement le signal. La plateforme d'expression est le domaine qui va permettre de réguler l'expression du gène grâce à un changement de conformation.

Les signaux reconnus par les riborégulateurs sont variés (Tableau 1).

Tableau 1. Liste des riborégulateurs connus.

Ligands	Gènes ¹	Fonctions	Références
Acides aminés et dérivés			
Lysine	<i>lysC</i>	Biosynthèse de la lysine	(Sudarsan et al., 2003)
Glycine	<i>gcvT</i>	Métabolisme de la glycine	(Mandal et al., 2004)
Glutamine	-	Métabolisme de l'azote	(Ames et Breaker, 2011)
Coenzymes			
AdoCbl ²	<i>btuB</i>	Transport de cobalamines	(Nahvi et al., 2002)
TPP ²	<i>thiC, thiM, tbpA</i>	Biosynthèse, transport de la TPP	(Winkler et al., 2002a)
FMN ²	<i>ribB</i>	Biosynthèse et transport du FMN	(Winkler et al., 2002b)
SAM ²	-	Biosynthèse (méthionine/cystéine/SAM)	(Wang et Breaker, 2008)
SAH ²	<i>ahcY, methH</i>	Recyclage SAM	(Wang et al., 2008)
Moco ²	<i>moaA</i>	Biosynthèse MoCo	(Regulski et al., 2008)
Tuco ²	-	-	
THF ²	<i>folT, folE, folC, folQ</i>	Transport et biosynthèse du THF	(Ames et al., 2010)
Bases azotées			
Adénine	<i>pbuE, yhdL</i>	Transport de purines	(Mandal et Breaker, 2004)
Guanine	<i>purE, xpt</i>	Homéostasie métabolique des purines	(Mandal et al., 2003)

Ligands	Gènes ¹	Fonctions	Références
2'dG ²	<i>guaA</i>	?	(Kim et al., 2007)
préQ1 ²	<i>queC</i>	Biosynthèse de la queuosine	(Roth et al., 2007)
c-di-GMP ²	<i>gbpA, tfoX</i>	Variées	(Sudarsan et al., 2008)
c-di-AMP ²	<i>ydaO</i>	Variées	(Nelson et al., 2013)
Ions			
Mg ²⁺	<i>mgtA, mgtE</i>	Transport du Mg ²⁺	(DeRose, 2007)
Mn ²⁺	<i>mntH, mntP</i>	Transport du Mn ²⁺	(Dambach et al., 2015)
F ⁻	<i>crcB, eriC</i>	Transport d'ions, variées	(Baker et al., 2012)
Paramètres physico-chimiques			
Température	<i>ibpA, hsp, rpoH, cspA</i>	Chocs thermiques	(Kortmann et Narberhaus, 2012)
pH	<i>alx</i>	Transporteur putatif	(Nechooshtan et al., 2009)
Autres			
GlcN6P ²	<i>glmS</i>	Biosynthèse de GlcN6P	(Winkler et al., 2004)
ARNt non chargé	<i>tyrS, glyQ</i>	Biosynthèse et import d'ARNt	(Henkin et Grundy, 2006)
Aminoglycoside	<i>aaC</i>	Résistance aux aminoglycosides	(Jia et al., 2013)

¹la liste des gènes régulés n'est pas exhaustive, ²signification des abréviations : AdoCbl : adénosylcobalamine, TPP : thiamine pyrophosphate, FMN : flavine mononucléotide, S-adenosylmethionine (SAM), S-adenosylhomocysteine (SAH), Moco : cofacteur molybdène, Tuco : cofacteur tungstène, THF : tétrahydrofolate, 2'dG : 2'-désoxyguanosine, préQ1 : précurseur queuosine 1 (7-aminométhyl-7-déazaguanine), c-di-GMP : diguanine monophosphate cyclique, c-di-AMP : diadénine monophosphate cyclique, glucosamine- 6-phosphate (GlcN6P)

Les riborégulateurs sont généralement conservés aussi bien au niveau de la séquence qu'au niveau de leur structure secondaire. Ces caractéristiques ont permis de prédire de nouveaux riborégulateurs potentiels grâce à des études bio-informatiques (Barrick et al., 2004; Weinberg et al., 2007). Certains de ces riborégulateurs ont été confirmés, ajoutant à la diversité des

riborégulateurs connus. Les autres, dont le ligand (ou le stimulus) n'est pas connu, sont appelés riborégulateurs orphelins. Des riborégulateurs ont également été décrits chez les eucaryotes. Ces derniers régulent l'expression génique en contrôlant l'épissage des ARNm (Li et Breaker, 2013).

Les riborégulateurs contrôlent l'expression du gène en aval au niveau de l'élongation de la transcription, de la stabilité du messager ou de l'initiation de la traduction. Les riborégulateurs sont ainsi divisés en trois catégories, selon le mécanisme de régulation utilisé : il y a les riborégulateurs transcriptionnels, les riborégulateurs qui contrôlent la stabilité de l'ARNm et les riborégulateurs traductionnels (Bastet et al., 2011).

1.2.4.1. Riborégulateurs transcriptionnels

Il en existe deux types : ceux qui dépendent de la formation d'un terminateur de transcription intrinsèque classique (c'est-à-dire une tige boucle riche en GC suivie d'une queue poly-U) et ceux qui dépendent de la présence de la protéine Rho sur l'ARNm (Bastet et al., 2011).

Le premier mécanisme est observé notamment pour les riborégulateurs *mgtE*, *mntH*, *btuB*, FMN, SAM, guanine et adénine chez *Bacillus subtilis* (Dambach et al., 2015; DeRose, 2007; Mandal et Breaker, 2004; Wang et Breaker, 2008; Winkler et al., 2002b). Excepté pour le riborégulateur adénine, la liaison du ligand à l'aptamère du riborégulateur induit un changement de conformation menant à la formation d'un terminateur de transcription intrinsèque classique (Figure 5). Dans le cas du riborégulateur adénine, la présence du ligand empêche la formation de ce terminateur et permet alors l'activation et non la répression du gène régulé. Ce type de terminateur n'est pas le seul chez *E. coli* et n'est sûrement pas majoritaire, mais il est très difficile de prédire d'autres types de terminateurs non-canoniques (Mitra et al., 2009).

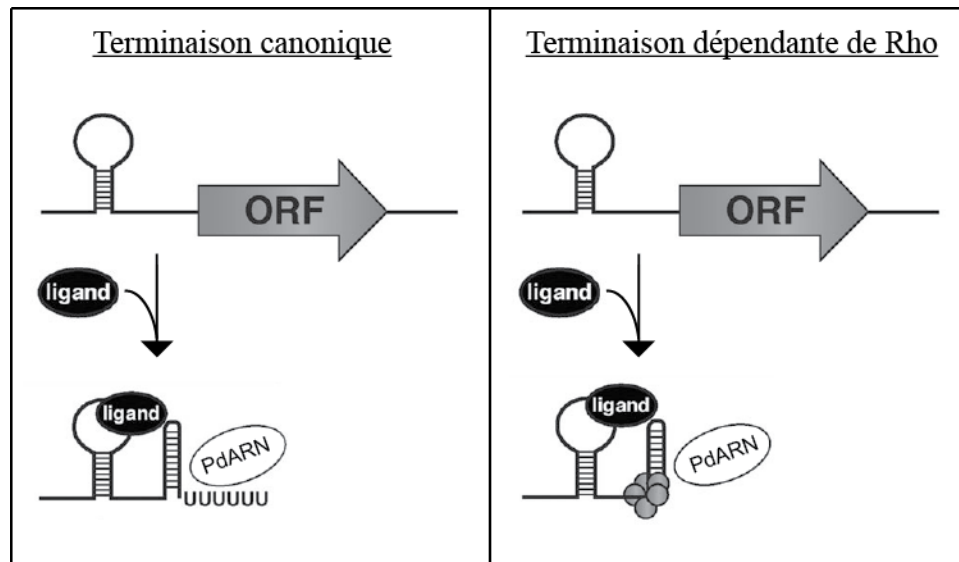


Figure 5. Mécanismes de régulation transcriptionnelle. Adapté de Bastet et al., 2011

Le deuxième mécanisme est utilisé par les riborégulateurs FMN chez *E. coli* et *mgtA* chez *Salmonella enterica* (Hollands et al., 2012) et possiblement par les riborégulateurs *lysC*, *thiC* et *thiM* chez *E. coli* (résultats non publiés, du laboratoire du Pr Daniel Lafontaine). Les séquences de ces riborégulateurs ne contiennent pas de terminateur intrinsèque classique, mais un mécanisme d'atténuation de la transcription dépendant de Rho est observé. Une étude récente a montré que la région 5' de l'ARNm *corA* de *S. enterica* peut adopter plusieurs structures et que dépendamment de la conformation adoptée la protéine Rho peut terminer la transcription de ce gène prématurément. Bien qu'aucun ligand n'ait été identifié pour l'instant, il est possible que l'expression de *corA* soit régulée par un riborégulateur dépendant de Rho (Kriner et Groisman, 2015).

1.2.4.2. Riborégulateurs contrôlant la stabilité de l'ARNm

Le riborégulateur *lysC* chez *E. coli* contrôle l'initiation de la traduction comme la plupart des riborégulateurs chez *E. coli* (cf. paragraphe sur les riborégulateurs traductionnels 1.2.4.3).

Cependant, une étude récente a montré qu'en plus du contrôle de la traduction, ce riborégulateur contrôle la stabilité de l'ARNm de manière ligand-dépendante (Caron et al., 2012). La liaison de la lysine à l'aptamère du riborégulateur entraîne un changement de conformation qui, en plus de séquestrer le RBS, libère un site reconnu et clivé par le dégradosome, menant à la dégradation du transcrit (Figure 6).

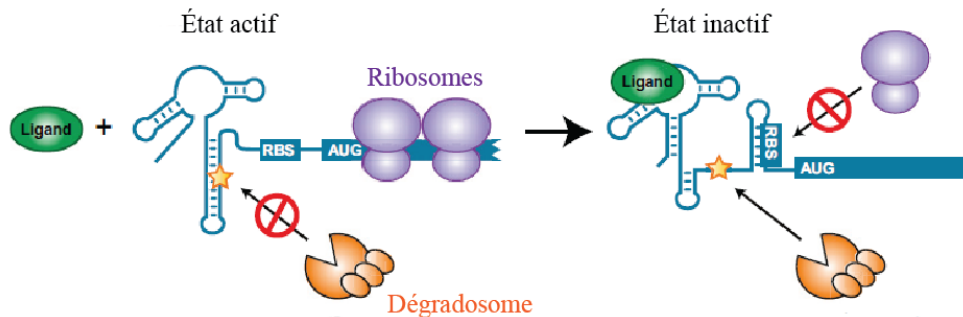


Figure 6. Mécanisme de régulation d'un riborégulateur dégradationnel. En absence de ligand, le riborégulateur adopte sa conformation active, séquestrant le site de clivage de la RNase E. Lorsque le ligand se lie, il induit un changement de conformation qui libère le site de clivage pour que le dégradosome puisse initier la dégradation de l'ARNm. Adapté de (Caron et al., 2012).

Le riborégulateur *lysC* est le seul riborégulateur connu adoptant ce mécanisme. Le riborégulateur *mgtA* chez *E. coli* pourrait néanmoins lui aussi utiliser la RNase E pour contrôler la stabilité de son transcrit (Spinelli et al., 2008).

Ce type de mécanisme pourrait donc s'étendre à d'autres riborégulateurs, et d'autres RNases pourraient être utilisées par des riborégulateurs pour réguler la quantité de transcrit. L'ARNm *corA*, chez *E. coli*, peut être dégradé par la RNase III (Lim et al., 2012). Étant donné que la région 5' non traduite de ce messenger chez *S. enterica* peut adopter plusieurs conformations, celle-ci pourrait contenir un riborégulateur contrôlant la stabilité de l'ARNm en recrutant la RNase III.

Un autre riborégulateur connu régule la stabilité de l'ARNm en s'autoclivant : le riborégulateur-ribozyme *glmS* (Ferré-D'Amaré et Scott, 2010). La liaison du glucosamine 6-phosphate au ribozyme induit un clivage séquence-spécifique.

1.2.4.3. Riborégulateurs traductionnels

Les riborégulateurs *lysC*, FMN, TPP et *btuB* chez *E. coli* contrôlent l'initiation de la traduction en modulant l'accessibilité du site de liaison des ribosomes (Figure 7).

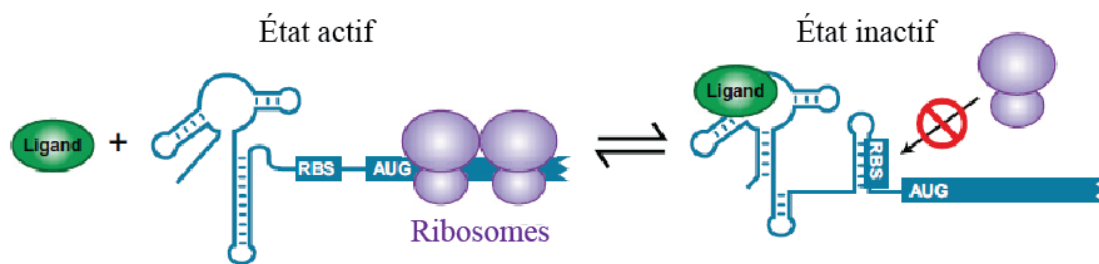


Figure 7. Mécanisme de régulation d'un riborégulateur traductionnel. En absence de ligand, le riborégulateur adopte sa conformation active, laissant le SD libre pour que les ribosomes traduisent activement l'ARNm. Lorsque le ligand se lie, il induit un changement de conformation qui mène à la séquestration du SD et inhibe ainsi la traduction. Adapté de Caron et al., 2012.

Les riborégulateurs qui détectent des changements de températures, appelés « ARN thermomètres » utilisent également le contrôle de la traduction pour réguler leurs gènes (Kortmann et Narberhaus, 2012). Il en est de même pour le régulateur pH (Nechooshtan et al., 2009).

En plus de ces quatre mécanismes, des études récentes ont permis la découverte de nouveaux mécanismes impliquant la traduction du riborégulateur (Park et al., 2010 ; Zhao et al., 2011) ou la liaison d'un petit ARN (résultats non publiés).

Une étude en cours dans le laboratoire révèle que le riborégulateur *btuB* chez *E. coli* est la cible de petits ARN régulateurs (pARN) (ou « small RNA»). Ces pARN contrôlent l'expression de *btuB*, indépendamment du ligand (CoB12) connu, en bloquant l'accès des ribosomes aux RBS et en induisant la dégradation de l'ARNm.

Les riborégulateurs peuvent donc utiliser des mécanismes et des signaux divers pour réguler l'expression génique, ce qui en fait des acteurs majeurs de l'adaptation aux variations de l'environnement.

1.2.5. LES PEPTIDES LEADERS

Certaines régions 5' des ARNm peuvent être traduites afin de réguler l'expression du gène situé en aval. Les cadres de lecture traduits sont généralement courts et les peptides produits sont appelés peptides leaders. Les régions 5' sont alors appelées leader, et pas « non-traduites » puisqu'elles contiennent un cadre de lecture pouvant être traduit. Les premiers peptides leaders ont été découverts dans les régions leaders des opérons tryptophane *trp* (Jackson et Yaetofsky, 1973) et histidine *his* (Kasai, 1974), puis dans celles d'autres opérons de biosynthèse d'acides aminés. La région leader de *trpE* chez *E. coli* contient un cadre de lecture de 14 codons, contenant deux codons tryptophane (trp) consécutifs, ce qui sort de l'ordinaire puisque le codon trp est plutôt rare, avec une occurrence d'environ un tous les 100 codons (Yanofsky, 1981). En cas de carence en trp, le ribosome pause au niveau des codons trp, libérant la région en aval (2) qui agit comme anti-terminateur en s'appariant avec la branche 5' du terminateur (3), permettant la transcription des gènes de biosynthèse du trp (Figure 8). À l'inverse, lorsqu'il y a assez de trp, la traduction du peptide leader se fait au complet, masquant l'anti-terminateur et permettant ainsi la formation du terminateur de transcription, réprimant ainsi l'expression des gènes de l'opéron. Dans le cas de l'opéron *his*, la régulation dépend de la disponibilité en ARNt chargés en histidine (his). En présence d'une faible

concentration de ceux-ci, le ribosome pause sur le cadre de lecture *hisL*, au niveau des 7 codons consécutifs *his* (sur 16 codons), empêchant la formation d'un terminateur de transcription canonique, en libérant la région anti-terminatrice (Alifano et al., 1996). Plus tard, des mécanismes similaires ont été décrits pour les opérons *thrA*, *leu*, *pheA* et *ilvG* responsables de la biosynthèse des acides aminés thréonine (*thr*), leucine (*leu*), phénylalanine (*phe*) et isoleucine (*ile*), et valine (*val*) respectivement (Gardner, 1979; Gemmill et al., 1979; Lawther et Hatfield, 1980; Zurawski et al., 1978). Les opérons *thrA* et *ilvG* présentent de plus la particularité de posséder plus qu'un type de codon surreprésenté : le peptide leader de *thrA* possède 4 *ile* en plus des 8 *thr*, sur 21 codons au total ; celui de *ilvG* contient 5 *ile*, 6 *val* et 4 *leu* sur 32 total, et est ainsi régulé par les ARNt^{ile}, ARNt^{val} et ARNt^{leu}. Le mécanisme de terminaison peut être dépendant de Rho (*trpA*, *pheA* et *ilvG*) ou non (*his*, *leu*) et dans le premier cas, le ribosome *calé* peut masquer le site *rut*.

L'opéron *trp* chez *Bacillus subtilis* ne contient pas de cadre de lecture dans sa région leader, mais utilise un mécanisme différent pour réguler la biosynthèse du *trp* en fonction de sa disponibilité. La région leader possède onze trinuécléotides G/U₁, A₂, G₃ reconnus par le complexe protéique TRAP (Tryptophan Regulated Attenuation Protein), composé de onze sous unités de la protéine, liées et stabilisées par le *trp* (Gollnick, 1994). La présence du complexe lié à la région leader de l'opéron permet à un terminateur intrinsèque de se former et à l'inverse lorsque le *trp* est limitant le complexe TRAP ne peut pas se lier à l'ARN, libérant la région anti-terminatrice menant à l'expression des gènes de biosynthèse du *trp*.

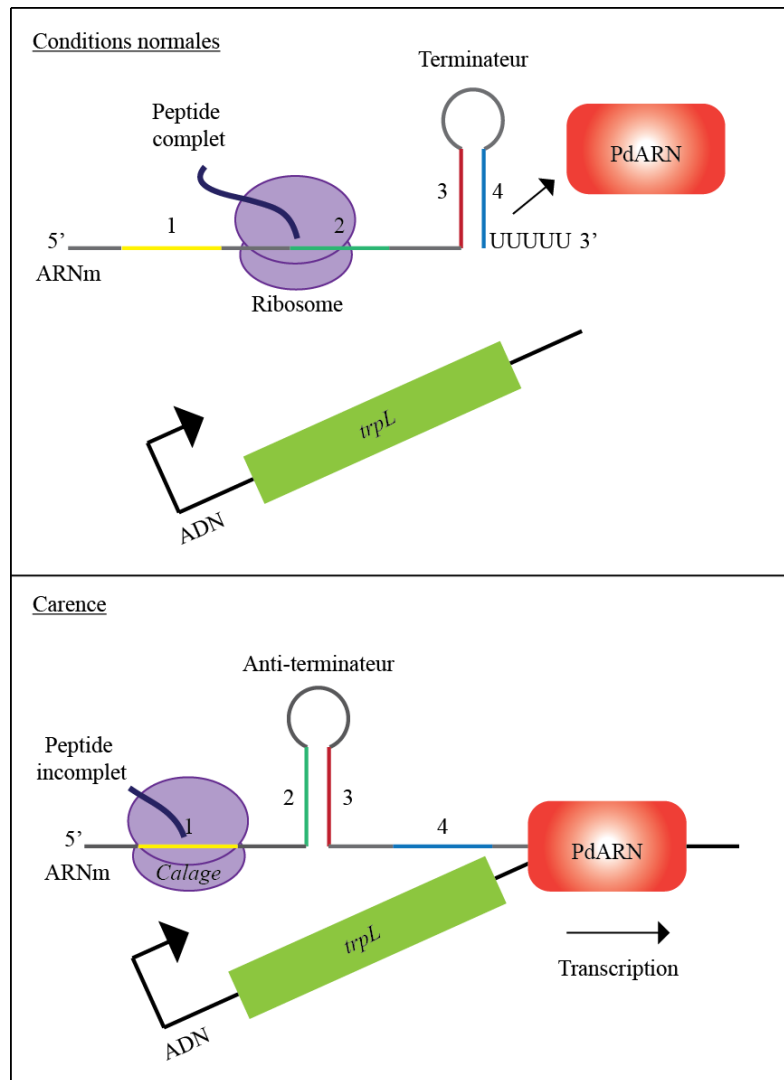


Figure 8. Mécanisme de régulation de l'opéron *trp* chez *E. coli*. En conditions normales, la traduction complète du peptide permet la formation d'un terminateur de transcription. En carence en tryptophane, le ribosome cale au niveau des codons *trp* de *trpL*, favorisant ainsi la formation d'une tige-boucle antiterminatrice. La plupart des peptides leaders connus utilisent un mécanisme similaire, impliquant le calage du ribosome en réponse à un signal cellulaire et modulant la formation du terminateur de transcription.

L'opéron *tna* qui code pour la tryptophanase, qui permet l'utilisation du *trp* comme source de carbone, d'azote et d'énergie (Hopkins et Cole, 1903), est également régulé par le *trp* chez *E. coli*. De manière similaire à l'opéron *trp*, la région leader contient un cadre de lecture de 21 codons, nommé *tnaC*. Ce peptide leader contient cependant un unique codon *trp* qui est

nécessaire à la régulation par le *trp*. À l'inverse de l'opéron *trp*, le ribosome ne cale pas au niveau du codon *trp* mais au niveau du codon de terminaison de la traduction du peptide (UGA) (Konan et Yanofsky, 1997). En absence de *trp*, le peptide est traduit complètement et le ribosome décroche de l'ARN, libérant un site *rut* qui va permettre à la protéine Rho d'induire la terminaison de la transcription, en rattrapant la PdARN pausée. En présence de *trp*, le peptide naissant induit le *calage* au codon stop, masquant ainsi le site *rut* pour que l'opéron soit transcrit entièrement. Le mécanisme par lequel le *trp* agit sur le ribosome pour induire son *calage* n'a été décrit que plus récemment (Bischoff et al., 2014). La présence de 2 molécules de *trp* dans le tunnel de sortie induit un changement de conformation du centre peptidyl-transférase qui empêche le facteur RF2 de reconnaître efficacement le codon stop afin de libérer le peptide pour terminer la traduction.

En plus des peptides leaders sensibles aux acides aminés, un autre groupe répond à différents antibiotiques. Les deux premiers peptides leaders contrôlant l'expression de gènes de résistance découverts sont retrouvés dans la région leader des gènes *cat* et *erm* (Dubnau et Nomura, 1984; Duvall et al., 1983). Dans le cas des gènes de résistance au chloramphénicol *cat* et *cmlA*, leur traduction est inhibée par une structure séquestratrice du SD. L'induction de ces gènes survient lorsqu'un ribosome cale sur le cadre de lecture de 9 codons présents en amont. La séquence du peptide dicte le site de *calage* mais nécessite la présence de chloramphénicol pour que le *calage* soit suffisamment fort pour empêcher la formation de la tige séquestratrice du SD (Lovett, 1996).

Dans le cas du gène *ermC* encodant un gène de résistance à l'érythromycine, un antibiotique de la classe des macrolides, un cadre de lecture de 19 codons nommé *ermCL* permet son expression en présence de l'antibiotique (Vazquez-Laslop et al., 2008). Le ribosome cale à l'ile 9 et est incapable de former le lien peptidique nécessaire à l'ajout de la serine en position 10 pour continuer l'élongation de la traduction du peptide. L'identité des 4 résidus en C-

terminal et la présence de l'antibiotique lié au ribosome dans le tunnel de sortie sont requises pour le *calage* et donc l'induction de l'expression de *ermC*. En absence, d'érythromycine, le peptide leader est traduit entièrement et une tige séquestratrice du SD empêche la traduction de *ermC*. En présence d'antibiotique, le *calage* du ribosome induit un changement de conformation libérant le SD, induisant ainsi la traduction de la cassette de résistance à l'antibiotique. Une étude plus récente (Gupta et al., 2013) a montré que les kétolides, qui n'induisent pas le *calage* comme l'érythromycine, peuvent tout de même induire l'expression de *ermC*. Cette induction est due à un décalage de phase (*frameshift*) de la traduction de *ermCL* qui permet au ribosome de dépasser le codon stop, ce qui mène également à un changement de conformation qui libère le SD de *ermC*.

Le transporteur à large spectre d'agents antibiotiques (*multidrug*) BmrC/BmrC chez *B. subtilis* possède également un peptide leader qui permet l'activation de l'expression du transporteur en présence de divers antibiotiques (Reilman et al., 2014). Un terminateur de transcription canonique situé à la fin du cadre de lecture du peptide *bmrB* serait à l'origine de la régulation observée avec les antibiotiques ciblant le ribosome. Un *calage* induit par la présence d'antibiotique et la présence de codons rares permettraient la formation de l'antitermineur.

Récemment, une étude a expliqué le contrôle du ratio substrat/enzyme de l'opéron *mccABCDE* (Zukher et al., 2014). Le gène *mccA* code pour le précurseur de la microcine C, les gènes *mccB*, *mccD* et *mccE* codent pour des enzymes nécessaires à sa maturation et *mccC* pour une pompe qui exporte ce peptide antimicrobien. Afin d'assurer une expression importante de *mccA* mais pas des autres gènes de l'opéron, la présence de ribosome sur *mccA* stabilise une structure terminatrice intrinsèque (mais non canonique) localisée entre *mccA* et *mccB*. Dans ce cas, le peptide leader a une fonction biologique contrairement aux autres peptides présentés ci-dessus.

Il existe d'autres systèmes de peptides leaders. Deux d'entre eux qui divergent de ceux présentés ci-dessus par le fait que le peptide naissant contient un signal d'intégration à la membrane, soit SecM et MifM. Le cadre de lecture *secM* est situé en amont du gène *secA* codant pour une ATPase, moteur du système de translocation des protéines à travers la membrane plasmique, et joue le rôle de moniteur de sécrétion. Chez *E. coli*, ce peptide leader est long (170 résidus) et contient une séquence de pause traductionnelle à la fin du cadre de lecture, induite par la chaîne peptidique naissante à l'intérieur du tunnel de sortie. La durée de la pause est contrôlée par la région N-terminale du peptide SecM naissant qui contient une séquence signal qui est reconnue par le système de sécrétion SecYEG. Lorsque celui-ci est fonctionnel, la translocation de SecM va permettre au ribosome de continuer la traduction (Butkus et al., 2003). Dans le cas contraire, le ribosome *calé* sur *secM* va permettre un changement de conformation de l'ARNm libérant le SD de *secA*, activant ainsi son expression. Le peptide leader MifM chez *B. subtilis* possède également une séquence de *calage* à la fin de son cadre de lecture et un segment transmembranaire N-terminal. Il sert de moniteur du système d'insertion des protéines à la membrane en régulant l'expression du gène *yidC* dont c'est le rôle (Chiba et al., 2009). Lorsque celui-ci est fonctionnel, l'insertion du peptide MifM naissant à la membrane force le ribosome *calé* à continuer la traduction. Dans le cas où celui-ci est défectueux, le *calage* prolongé du ribosome induit un changement de conformation de l'ARNm, libérant le SD de *yidC*, activant ainsi son expression.

Récemment, des études ont montré qu'un riborégulateur pouvait également utiliser la traduction d'un peptide leader pour réguler l'expression génique : le riborégulateur *mgtA* chez *S. enterica* (Gall et al., 2016; Park et al., 2010; Zhao et al., 2011) (le mécanisme de régulation est décrit dans le paragraphe 3 sur le riborégulateur *mgtA*).

1.2.6. LA RÉGULATION PAR LE Mg^{2+}

Comme mentionné plus haut, le magnésium est essentiel à un grand nombre de processus biologiques. Contrôler la concentration intracellulaire et donc son transport cellulaire est donc crucial pour la survie. Cela se traduit par une régulation à différents niveaux de l'expression des transporteurs, mais également à une cascade de régulation en réponse aux variations de concentration en Mg^{2+} . Jusqu'à présent, tous les gènes connus pour être régulés par le Mg^{2+} le sont au moins au niveau de leur promoteur *via* le système à deux composants PhoP/PhoQ.

Le système à deux composants PhoP/PhoQ contrôle plusieurs gènes de virulence chez *S. enterica*, mais est également un senseur de Mg^{2+} extracellulaire, notamment chez *E. coli* (Figure 9). La protéine PhoQ est le senseur du Mg^{2+} qui change de conformation en présence de Mg^{2+} périplasmique. PhoP est la protéine régulatrice du système qui est nécessaire pour l'expression de nombreux gènes. Elle est phosphorylée par la kinase PhoQ lorsque celle-ci subit un changement de conformation en réponse à une faible concentration de Mg^{2+} par exemple (Garcia Vescovi et al., 1997). La forme phosphorylée de PhoP est capable d'activer un ensemble de gènes variant d'une espèce bactérienne à une autre (Perez et al., 2009). En plus d'une carence en Mg^{2+} , le système PhoPQ est activé par un pH acide (Prost et al., 2007) et par une carence en Ca^{2+} , et est réprimé par des peptides antimicrobiens cationiques (CAMP : cationic antimicrobial peptides) (Bader et al., 2003). Bien que la plupart des gènes cibles soient activés par la forme phosphorylée de PhoP, certains gènes peuvent être réprimés tels que le gène *treR* (Yamamoto et al., 2002).

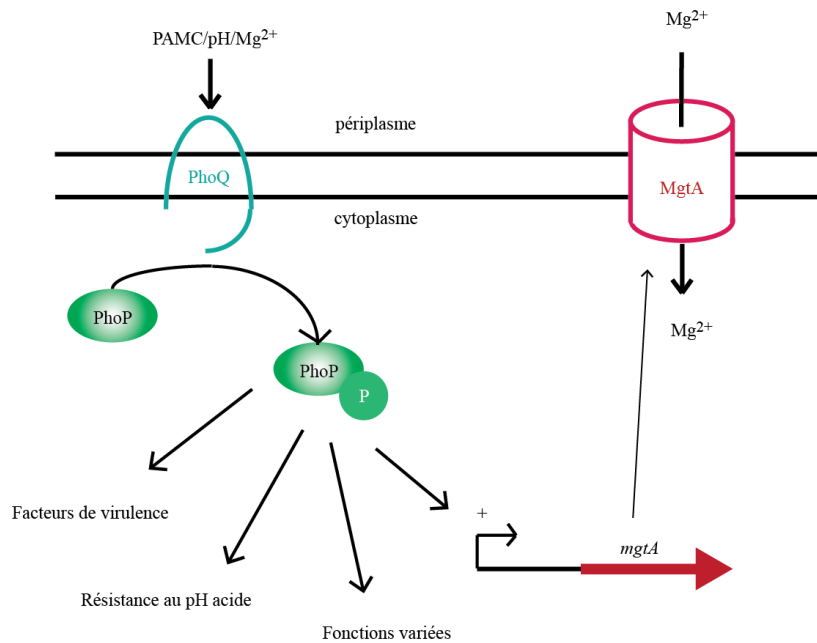


Figure 9. Régulation par le système à deux composants PhoQ/PhoP. Le senseur PhoQ détecte un changement de concentration en Mg^{2+} ou de pH ou la présence de peptides antimicrobiens cationiques (PAMC). Le changement de conformation induit par ses signaux lui permet de réguler le niveau de phosphorylation du régulateur PhoP. Celui-ci lorsqu'il est phosphorylé active la transcription de nombreux gènes, impliqués dans le transport de Mg^{2+} (*mgtA*), la virulence, la résistance au pH acide, et d'autres fonctions diverses.

Outre le système PhoQ/PhoP, un autre mécanisme de régulation de l'expression génique par le Mg^{2+} existe au niveau post-transcriptionnel : il s'agit de riborégulateurs Mg^{2+} . Il en existe deux différents, localisés en amont des gènes *mgtA* et *mgtE*. Le riborégulateur *mgtA*, présent notamment chez *S. enterica* et *E. coli*, sera décrit dans le paragraphe 3. Le riborégulateur *mgtE* a été découvert chez *B. subtilis* et est absent chez *S. enterica* et *E. coli* (Dann III et al., 2007). Ce riborégulateur dont l'aptamère est nommé M-box régule au niveau transcriptionnel grâce à un terminateur canonique de la transcription.

1.3. LE RIBORÉGULATEUR *MGT A* CHEZ *E. COLI*

Le gène *mgtA* code pour un transporteur de Mg^{2+} et son expression est régulée par différents facteurs. Au niveau de son promoteur, deux facteurs de transcription peuvent intervenir : Rob et PhoP. Ce premier active la transcription de *mgtA* chez *S. enterica* (Barchiesi et al., 2008). Ce facteur de transcription est associé à un phénotype de résistance à plusieurs agents antibiotiques, aux métaux lourds, à l'ion superoxyde et aux solvants chez *E. coli*. Le promoteur proximal de *mgtA* répond à la concentration extracellulaire en magnésium *via* le système à deux composants PhoQ/PhoP. Lorsque la concentration en magnésium est faible, le senseur PhoQ phosphoryle son régulateur PhoP qui est capable de se lier au promoteur de *mgtA* seulement dans sa forme phosphorylée. Ainsi une faible concentration en magnésium active l'initiation de la transcription de *mgtA* (Minagawa et al., 2003).

En plus de la régulation au niveau du promoteur, un riborégulateur a été découvert dans le 5' de *mgtA*. Celui-ci a été découvert et étudié principalement chez *S. enterica* bien qu'il soit conservé chez plusieurs espèces.

1.3.1. LE RIBORÉGULATEUR CHEZ *S. ENTERICA*

Une étude récente (Cromie et al., 2006) a montré un deuxième niveau de régulation dépendante du magnésium en identifiant un riborégulateur dans la région 5' non traduite de *mgtA* chez *S. enterica*. Les résultats de l'étude suggèrent que le riborégulateur est dans sa conformation active lorsque la concentration extracellulaire est inférieure à 50 μM et dans sa conformation inactive à partir de cette concentration. La question de la spécificité du riborégulateur au magnésium a été adressée en comparant l'effet du Ni^{2+} , du Zn^{2+} , du Ca^{2+} et du Mn^{2+} à celui du Mg^{2+} pour une même concentration (25 μM). Les ions Ca^{2+} et Mn^{2+} ne semblent avoir aucun effet. Les ions Ni^{2+} et Zn^{2+} ont un effet, mais moins prononcé que celui

du Mg^{2+} . Les auteurs de l'étude ont également identifié une région pouvant adopter deux conformations différentes en fonction de la concentration en magnésium (Figure 10). Les tiges boucles A et B seraient formées à de fortes concentrations en magnésium et la tige boucle C à de faibles concentrations.

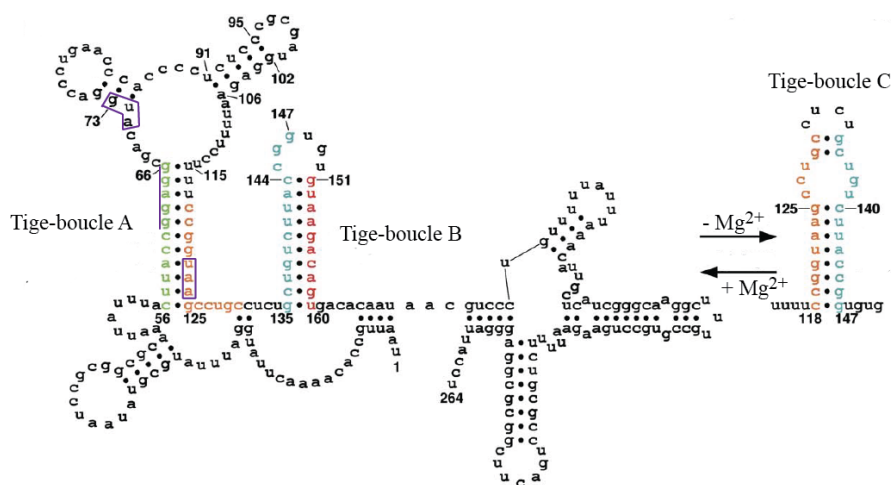


Figure 10. Structures du riborégulateur *mgtA* chez *S. enterica*. La région 3' de la tige-boucle A (en orange) s'apparie avec la région 5' de la tige-boucle B (en bleu) pour former la tige-boucle alternative C en faible Mg^{2+} . L'AUG et le codon stop du peptide leader *mgtL* sont entourés en violet. Son SD est indiqué par une ligne violette. Adapté de Cromie et al., 2006.

Le mécanisme proposé dans cette étude est un arrêt de transcription indépendant de Rho bien qu'aucun terminateur classique n'ait été identifié. Une étude du même groupe s'est intéressée à l'intérêt d'une double régulation promoteur-riborégulateur (Cromie et Groisman, 2010). Trois gammes distinctes de concentration de magnésium : à de très faibles concentrations (inférieures à 50 μM), le promoteur et le riborégulateur sont actifs, à des concentrations moyennes (entre 50 μM et 1 mM environ), le promoteur est actif, mais le riborégulateur inactif, et à de fortes concentrations (supérieures à 1 mM), ni le promoteur ni le riborégulateur ne sont actifs. Le riborégulateur étant produit à l'intérieur de la bactérie, il est admis qu'il

répond à la concentration intracellulaire en magnésium. Il est également admis qu'en faisant varier la concentration extracellulaire, la concentration intracellulaire va également varier. Pour vérifier ces hypothèses, les auteurs ont surexprimé *corA*, codant pour un autre transporteur de magnésium, dans le but d'augmenter la concentration cytoplasmique de magnésium. Cette surexpression a pour effet d'inhiber l'expression de *mgtA* via le riborégulateur. Ces résultats suggèrent donc que le riborégulateur répond bien à des concentrations intracellulaires et qu'en modulant la concentration extracellulaire, la concentration intracellulaire est également modulée.

Récemment, deux études de la même équipe ont montré que la protéine Rho était nécessaire à la régulation exercée par le riborégulateur *in vivo* (Hollands et al., 2012, 2014). Celles-ci montrent de plus que Rho favorise une terminaison de transcription à de fortes concentrations en magnésium *in vitro*. Ils ont identifié un site *rut* potentiel. Un site de pause particulièrement long (avec une demi-vie de 174 secondes dans leurs conditions *in vitro*) a été détecté à l'intérieur du riborégulateur. Cette pause de l'ARN polymérase est plus longue à de fortes concentrations de magnésium et nécessite la formation de structures formées dans ces conditions de Mg^{2+} , suggérant son importance pour le fonctionnement du riborégulateur. Ils montrent de plus que cette pause est nécessaire à la terminaison dépendante de Rho et qu'elle se situe au niveau de l'arrêt de la transcription.

Une étude plus récente (Spinelli et al., 2008), montre qu'en plus du mécanisme d'atténuation de la transcription, la région 5' non traduite de *mgtA* peut être clivée par la RNase E en présence de fortes concentrations en Mg^{2+} .

En plus du mécanisme dépendant de Rho, le riborégulateur pourrait utiliser la traduction d'un peptide leader pour réguler l'expression de *mgtA* (Gall et al., 2016; Park et al., 2010; Zhao et al., 2011). En effet, un petit cadre de lecture (17 codons chez *E. coli*), nommé *mgtL*, est

retrouvé dans le riborégulateur de plusieurs espèces. Il peut être traduit, mais le peptide en lui-même ne semble pas avoir d'effet en *trans* sur l'expression de *mgtA*. Les deux études montrent que la traduction du peptide favorise l'arrêt de transcription Rho-dépendant. Cependant, les conclusions de ces études sont en désaccord quant au signal contrôlant sa traduction. Selon Park et al. (2010), la traduction du peptide est contrôlée par la disponibilité de la proline, et ce de manière indépendante du Mg^{2+} . À l'inverse, Zhao et al. (2011), suggèrent que la traduction du peptide est contrôlée par la concentration en magnésium. En effet, ils ne détectent le peptide qu'à fortes concentrations en magnésium et observent une modification de l'accessibilité du SD de *mgtL* en fonction du magnésium *in vitro*. L'étude la plus récente par Gall et al. (2016), suggère que l'expression de *mgtA* n'est pas régulée par un riborégulateur, mais uniquement par le peptide leader *mgtL* dont l'efficacité de traduction est affectée par le Mg^{2+} . Malgré leur désaccord sur le mécanisme de régulation du peptide leader, toutes les équipes démontrent l'importance de la traduction du peptide sur l'expression de *mgtA*.

1.3.2. LE RIBORÉGULATEUR MGTA CHEZ *YERSINIA ENTEROCOLITICA*

Le ou les sites de liaison du Mg^{2+} au riborégulateur et le mécanisme par lequel le changement de conformation a lieu sont encore inconnus. Cependant, une étude du riborégulateur chez *Y. enterocolitica* (Korth et Sigel, 2012) suggère que la tige-boucle C, formée en présence de faibles concentrations en Mg^{2+} , a une très forte affinité pour le magnésium. Cette région conservée et riche en AU pourrait permettre la détection de concentrations élevées en Mg^{2+} et permettre le changement de conformation.

1.3.3. LE RIBORÉGULATEURS MGTA CHEZ *E. COLI*

Chez *E. coli*, la région 5' non traduite de *mgtA* est longue et similaire en séquence et en structure à celles de *S. enterica* et *Y. enterocolitica*, suggérant la présence d'un riborégulateur

fonctionnel. Bien qu'aucune étude sur ce riborégulateur n'ait encore été publiée, on peut remarquer qu'en plus de pouvoir adopter des structures similaires à celles observées chez *S. enterica*, le riborégulateur chez *E. coli* contient un cadre de lecture similaire en longueur et en séquence et dans la même région du riborégulateur. De plus, une étude récente sur les sites de pauses chez *E. coli* (Larson et al., 2014) suggère la présence d'un site de pause au C214, c'est-à-dire dans la même région que celle où se trouve le site de pause chez *S. enterica*. Toutes ces observations suggèrent que le riborégulateur chez *E. coli* pourrait être fonctionnel et utiliser un mécanisme de régulation similaire à celui de *S. enterica*, qui est un organisme phylogénétiquement proche. Enfin, une étude récente a mis en évidence l'implication d'une petite protéine, nommée MgtS (précédemment YneM) dans la régulation post-traductionnelle de l'expression de *mgtA*. En effet, cette protéine dont l'expression est régulée par le Mg^{2+} via PhoP, s'associe à MgtA pour moduler sa capacité à importer du Mg^{2+} (Wang et al., 2017).

1.4. OBJECTIFS DU PROJET DE RECHERCHE

L'homéostasie du Mg^{2+} , bien que vitale pour la cellule, est encore mal comprise, aussi bien au niveau de l'identification de ses transporteurs que de la régulation des gènes par le Mg^{2+} . La régulation de cette homéostasie est complexe et cette complexité est illustrée par la régulation du gène *mgtA*. En effet, ce gène codant pour un transporteur de Mg^{2+} possède plusieurs niveaux de régulation dont les signaux et mécanismes restent sujets à discussion. Bien qu'étudié chez *S. enterica*, peu de littérature existe au sujet du même gène chez *E. coli*. C'est d'ailleurs le seul riborégulateur de *E. coli* qui n'a pas été étudié à ce jour.

Le premier objectif de cette thèse est donc l'étude du riborégulateur *mgtA* chez *E. coli*. Son existence sera vérifiée en montrant que celui-ci est capable de réguler l'expression de *mgtA* en fonction de la disponibilité en Mg^{2+} . L'étude de ses mécanismes de régulation permettra par la suite de le comparer à celui de *S. enterica*. L'objectif sera alors de déterminer l'importance

relative de chacun de ses mécanismes. Plus précisément, l'étude se concentrera à déterminer si le riborégulateur et le peptide leader sont deux mécanismes indépendants ou si, comme le suggèrent Gall *et al.* (2016) le peptide leader est le seul mécanisme. Le rôle de MgtS sera également évalué pour vérifier son effet sur le riborégulateur.

L'effet du Mg^{2+} sur l'expressome n'a jusqu'à présent été étudié que *via* le système à deux composants PhoQ/PhoP. La découverte du riborégulateur *mgtA*, qui régule l'expression de manière indépendante de PhoP ouvre la possibilité d'autres voies de régulation par le Mg^{2+} . Un deuxième objectif sera alors de vérifier si de telles régulations existent afin de mieux comprendre l'effet d'un changement en Mg^{2+} sur le métabolisme bactérien.

CHAPITRE 2 :

MATÉRIEL ET MÉTHODES

2.1. SOUCHES, PLASMIDES ET OLIGOS

Les souches, plasmides et oligonucléotides utilisés dans cette thèse sont listés dans les tableaux 2 (souches et plasmides) et 3 (oligonucléotides). Le tableau 4 décrit les stratégies de polymérisation en chaîne (*Polymerisation Chain Reaction*, PCR) utilisées.

Tableau 2. Liste des souches.

Nom	Description	Provenance
<i>Escherichia coli</i>		
EM1055	MG1655 $\Delta lacZ$ X174	(Massé et Gottesman, 2002)
EM1377	EM1055 <i>rne</i> -131 <i>zce</i> -726::Tn10	(Massé et al., 2003)
EM1277	EM1055 <i>rne</i> -3071(Ts) <i>zce</i> -726::Tn10	(Massé et al., 2003)
MG22	EM1055 + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (8 codons)	Cette étude
MG24	<i>rne</i> -131 + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (8 codons)	Cette étude
MG26	<i>rne</i> -TS + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (8 codons)	Cette étude
MG29	EM1055 + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (8 codons) Δ UCUU	Cette étude
MG30	EM1055 + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (123 premiers nt)	Cette étude
MG31	EM1055 + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (163 premiers nt)	Cette étude
MG32	EM1055 + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (193 premiers nt)	Cette étude
MG33	EM1055 + <i>PlacUV5-mgtA'</i> - <i>lacZ</i> (243 premiers nt)	Cette étude
MG34	EM1055 + <i>PlacUV5-lacZ</i>	Cette étude

Nom	Description	Provenance
MG51	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (+163, 8 codons)	Cette étude
MG55	MG22 + <i>psk260</i>	Cette étude
MG56	MG22 + <i>psk262c</i>	Cette étude
MG66	EM1055 <i>rnhA::kan</i>	Cette étude
MG67	EM1055 <i>rfaH::kan</i>	Cette étude
MG70	MG22 <i>rnhA::kan</i>	Cette étude
MG71	MG22 <i>rfaH::kan</i>	Cette étude
MG81	EM1055 + <i>p_{mgtA}-lacZ</i>	Cette étude
MG82	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (8 codons) mutSD	Cette étude
MG83	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (8 codons) AGG	Cette étude
MG84	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (8 codons) STOP	Cette étude
MG85	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (8 codons) GUG	Cette étude
MG86	EM1055 $\Delta p_{mgtA} :: PlacUV5$	Cette étude
MG112	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (+133, 8 codons)	Cette étude
MG160	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (8 codons) mut4U	Cette étude
MG161	EM1055 + <i>PlacUV5-thiM-mgtL-mgtA'-lacZ</i> (8 codons)	Cette étude
MG162	EM1055 + <i>PlacUV5-thiM'(17codons) -mgtA'-lacZ</i> (8 codons)	Cette étude
MG172	EM1055 + <i>PlacUV5-thiM-mgtL-lacZ</i>	Cette étude
MG182	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (8 codons) C214G	Cette étude
MG186	EM1055 + <i>PlacUV5-mgtA'-lacZ</i> (8 codons) de <i>S. enterica</i>	Cette étude
MG203	EM1055 + <i>PlacUV5-MgtA'-LacZ</i> (8 codons)	Cette étude
MG205	EM1055 + <i>PlacUV5-MgtL'-LacZ</i> (AUG)	Cette étude
MG206	EM1055 + <i>PlacUV5-MgtL'-LacZ</i> (GUG)	Cette étude
MG224	NM580 + <i>p_{endogène}-OmpX'-LacZ</i> (11 codons)	Cette étude
MG227	NM580 + <i>p_{endogène}-YnfB-LacZ</i> (8 codons)	Cette étude

Nom	Description	Provenance
MG231	EM1055 + <i>PlacUV5-thiM-mgtL-mgtA'-lacZ</i> (8 codons) de <i>S.</i>	Cette étude
MG240	MG83 + pRS1551	Cette étude
MG241	MG83 + pMgtL	Cette étude
MG261	EM1055 + <i>PlacUV5-corA'-lacZ</i> (8 codons)	Cette étude
MG270	EM1055 + <i>PlacUV5-thiM-calage-mgtA'-lacZ</i> (8 codons)	Cette étude
MG281	EM1055 + p _{endogène} - <i>corA'-lacZ</i> (8 codons)	Cette étude
MG284	EM1055 + <i>PlacUV5-thiM-corL-corA'-lacZ</i> (8 codons)	Cette étude
MG286	EM1055 + p _{prox} - <i>ybeX'-lacZ</i> (3 codons)	Cette étude
MG291	EM1055 <i>yneM</i> -	Cette étude
MG293	MG161 <i>yneM</i> -	Cette étude
NM580	MG1655, <i>lacIQ</i> -FRT scar, DlacZp::T1T2-ZeoR-kanR-pBAD- <i>ccdB</i> ; mini-lRed::tetR.	Laboratoire Massé
<i>Salmonella enterica</i>		
14028s	<i>S.enterica</i> sérovar Typhimurium WT	Laboratoire Massé
MG310	WT + pMLASe	Cette étude
MG311	WT + pMLAEc	Cette étude
Plasmides		
pRS1551	Plasmide pour fusions traductionnelles	(Simons et al., 1987)
pFRA	Plasmide pour fusions transcriptionnelles	(Repoila et Gottesman, 2001)
pMgtL	pRS1551 + région 5' de <i>mgtA</i> (promoteur <i>PlacO</i> , EcoRI+BamHI)	Cette étude
psk260	Plasmide de surexpression de la RNase H	(Drolet et al., 1995)
psk262c	Plasmide contrôle pour psk260	(Drolet et al., 1995)
pXG10	Plasmide pour fusions <i>gfp</i>	(Urban et Vogel, 2007)
pMLASe	pXG10 + <i>PlacUV5-thiM-mgtL-mgtA'-lacZ</i> (8 codons) de <i>S. enterica</i>	Cette étude

Nom	Description	Provenance
pMLAEc	pXG10 + PlacUV5- <i>thiM-mgtL-mgtA'</i> - <i>lacZ</i> (8 codons) de <i>E. coli</i>	Cette étude

Tableau 3. Tableau des amorces.

Nom	Séquence
AC687	AGTTTGCTCCTGCC
EM111	GAGGTACCAGCGCGGTTTGATC
EM112	TTAATATATTGATATTTATATCATTTTACGTTTCTCGTTC
EM113	ACTCGTCGCGAACCGCTTTC
EM194	GCCATAAACTGCCAGGAATTGG
EM195	CGGGCCTCTTCGCTA
MG201	TAACCTGGCCATTCTGAAAAAAGCGCCGCAAATAATAAG
MG202	TTTTCAGAATGGCCAGGTTATATGGAACCTGATCCCAC
MG205	ACGCCGTGAACGGGCTTACCGGAAAAGTTTTAATC
MG206	GGTAAGCCCGTTACGCGCTTACCGGATG
MG209	ACCGGAGGTTATGTGGAACCTGATCCC
MG210	GCGTGGGATCAGGTTCCACATAACCTC
MG213	ACCGGAGGTTATAGGGAACCTGATCCC
MG214	GCGTGGGATCAGGTTCCCTATAACCTC
MG215	TCGACGGAGATAAAAACTTTTCCGGTAAGC
MG216	GCTTACCGGAAAAGTTTTATCTCCGTCGAG
MG388	CGCGGATCCCCGGGTAAAAATTTCTTTAAACATAAGG
MG536	CCCCGGAATTCGGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATG TGTGGAATAGATGCTACGAATATTATTGGATTC
MG595	CGGGTAGCAAAACAGATCGAAGAAGGGGTGAATCGCAGGCGGGCACCCCAGG CTTTACAC
MG596	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAAAACGACTTCCACACATTATA CGAGCCGGAAG

Nom	Séquence
MG990	CGCGGATCCCTTACCGGAAAAGTTTAAATCTCC
MG994	CGCGGATCCTCACGGCCTTACGCATCC
MG997	CGCGGATCCGTAATAAAGCTGAGCAGGGACGTTAAAAC
MG998	CGCGGATCCTCTGAAACAGTACAGAAATGTCTTC
MG1055	CGCGGATCCTTCCACACATTATACGAGCCGGAAGC
MG1056	CCCCGGAATTCGGGCACCCCAGGCTTTACACTTTATG
MG1104	CCCCGGAATTCGGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATG TGTGGAATAATTGCCACAAAATTATGG
MG1105	CGCGGATCCCTGGCGGGTAATGATTTTTAG
MG1557	CCCCGGAATTCATAAATGTGAGCGGATAACATTGACATTGTGAGCGGATAACAA GATACTTAGATGCTACGAATATTATTGGATTC
MG1558	CGCGGATCCAAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTAAAGGAGTCCC TCCGCACTGTC
MG1601	AGCCTAACGTTACCCGTGCATTCAGCAATGGGTAAAGTCTGTGTAGGCTGGAGC TGCTTC
MG1602	TTCCACACATTATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCCCAT ATGAATATCCTCCTTAG
MG1603	TAGATGCTACGAATATTATTGGATTCTCCTTATTATTTGCTTCCACACATTATAC GAGCC
MG1604	AACGTTACCCGTGCATTCAG
MG1662	AAGGCTTCGCCAGGCCTGAAGAC
MG1663	CTTCAGGCCTGGCGAAGCCTTAC
MG1979	CCCCGGAATTCGGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATG TGTGGAAGCCCGTCTTTTCACGGCGTTAC
MG1970	CCCCGGAATTCGGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATG TGTGGAAACGTTTTAACGTCCCTGCTCAG
MG2010	CCCCGGAATTCTCAGCCGATTTTGATCCTG
MG2011	CGCGGATCCTGGCGATATAATACCTGCTGACAAC
MG2187	CCCCGGAATTCGGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATG TGTGGAATCACGGCGTTACCGGATGCG

Nom	Séquence
MG2552	CGCGGATCCCCGGAAGTTTTAATCTCCGTCGAGGGAGAG
MG2600	CGTAAGGCCGTGACGAAAAACGTCCCTGC
MG2601	AGCAGGGACGTTTTTTCGTCACGGCCTTAC
MG2604	GTTATGGCAGGAGCAAACCTATGGAACCTGATCCCACGCCTCTC
MG2681	GTTATGGCAGGAGCAAACCTATGGACCCTGAACCCACCCCTCTC
MG2685	CGCGGATCCATATAACCTCCGGTAAGTGAAAAAAG
MG2686	CGCGGATCCCCGGAAGTTTTAATCTCCGTCGAGGGAGAGGGCGTGGGATCAG
MG2687	CGCGGATCCCGGGTAAAAATTTCTTTAAACATAAGGAGTCCCTCCGCACTG
MG2779	CCCCGGAATTCGGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATG TGTGGAAGATCACTCTAAGAGGACATTCGCCTTGGAC
MG2780	CGCGGATCCTCCAGTTGAAATGCGCTCAGCATGAC
MG2840	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAAAACGACTCGTTTGCTGAGAG TGATTTTC
MG2841	ATACTATGCCGATATACTATGCCGATGATTAATTGTCAACCCAGACAAACAGCG CCAGTG
MG2859	ATACTATGCCGATATACTATGCCGATGATTAATTGTCAACACAACAGCGGTAAC CGTTGC
MG2861	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGCTGAAA GACATG
MG2862	ACTGGTCGAGTCAGATTTCTGCTCGAGCATGGCCCCCAATCGCATGCATCTAA GCCCCGTCTTTTCACGG
MG2863	GAGCAGAAATCTGACTCGACCAGTCTCGACGCTCATCATAGTTTGCTCCTGCCA TAACG
MG2899	CCCCGATGCATCTGCGATTTATCATCGCAACC
MG2904	GCTAGCGCGCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTTTACT GGCGGGTAATGATTTTTAGCATAGG
MG2906	GCTAGCGCGCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTTTACC GGGTAAAAATTTCTTTAAACATAAGGAGTCCC
MG3012	GACACACCCAGTAGATACTGGCTCACTATCC
MG3013	CAGTATCTACTGGGTGTGTCCATAGTTTGCTCCTGCCATAACGTGAAGAAGC

Nom	Séquence
MG3019	CCTCTGCTAATATCTTTCTCATTGAGATGAAAATTAAGGGTGTAGGCTGGAGCT GCTTC
MG3020	TAACAAGGGGCCGGTTAGGTGAGGGATTATCTCCGTTACATATGAATATCCTC CTTAG
MG3023	CCCCGGAATTCACAAACCGAGCGGCGCATAG
MG3024	CCCCGGAATTCTCTGCCGTCAGGTGGTTGAG
MG3027	CGCGGATCCTCGCTCATGGCGTCGTATGG

Tableau 4. Stratégies de PCR avec les amorces utilisées.

Constructions	Stratégie et Amorces
Souches	
MG22	MG536 + MG388
MG29	MG536+MG205 (fragment 5') et MG206+MG388 (fragment 3'), élongation puis MG536+MG388
MG30	MG536+MG990
MG31	MG536+MG994
MG32	MG536+MG997
MG33	MG536+MG998
MG34	MG595+MG596
MG51	MG1970+MG388
MG81	MG2010+MG2011
MG82	MG536+MG201 (fragment 5') et MG202+MG388 (fragment 3'), élongation puis MG536+MG388
MG83	MG536+MG214 (fragment 5') et MG213+MG388 (fragment 3'), élongation puis MG536+MG388
MG84	MG536+MG216 (fragment 5') et MG215+MG388 (fragment 3'), élongation puis MG536+MG388
MG85	MG536+MG210 (fragment 5') et MG209+MG388 (fragment 3'), élongation puis MG536+MG388

Constructions	Stratégie et Amorces
MG86	MG1601+MG1602 puis MG1601+MG1603 (PCR nichée)
MG112	MG2187+MG388
MG160	MG536+MG2601 (fragment 5') et MG2600+MG388 (fragment 3'), élongation puis MG536+MG388
MG161	MG2620+AC687 (fragment 5') et MG2604+MG388 (fragment 3'), élongation puis MG2620+MG388
MG162	MG2620+MG2614 (fragment 5') et MG1969+MG388 (fragment 3'), élongation puis MG2620+MG388
MG172	MG2620+MG2552 sur souche MG161
MG182	MG536+MG1663 (fragment 5') et MG1662+MG388 (fragment 3'), élongation puis
MG186	MG1104+MG1105
MG203	MG536+MG2687
MG205	MG536+MG2685
MG206	MG536+MG210 (fragment 5') et MG209+MG2686 (fragment 3'), élongation puis MG536+MG2686
MG224	MG2859+MG2861
MG227	MG2840+MG2841
MG231	MG2620+AC687 (fragment 5') et MG2681+MG1105 (fragment 3'), élongation puis MG2620+MG1105
MG261	MG2779+MG2780
MG270	MG2620+MG2863 (fragment 5') et MG2862+MG388 (fragment 3'), élongation puis MG2620+MG388
MG281	MG3023+MG2780
MG284	MG2779+MG3013 (fragment 5') et MG3012+MG2780 (fragment 3'), élongation puis MG2779+MG2780
MG286	MG3024+MG3027
MG291	MG3019+MG3020
Plasmides	
pMgtL	MG1557+MG1558
pMLASe	MG2899+MG2904
pMLAEc	MG2899+MG2906

L'ADN génomique utilisé pour les réactions de PCR a été isolé à partir de la souche EM1055 par la méthode d'extraction au phénol (Wright et al., 2017).

Les fusions *lacZ* ont été construites et insérées au niveau chromosomique comme décrit précédemment (Simons et al., 1987), en utilisant le plasmide pRS1551 pour les fusions traductionnelles et le pFRA pour les fusions transcriptionnelles. La région à fusionner a été amplifiée par PCR en ajoutant *PlacUV5* si nécessaire et les sites de restrictions EcoRI en 5' et BamHI en 3'. Après digestion avec les enzymes de restrictions de New England Biolabs (NEB), les inserts sont ligués dans le plasmide désiré avec la T4 DNA ligase de NEB selon les spécifications du fournisseur. Le mélange de ligation est par la suite transformé dans une souche DH5 α et les cellules transformées sont étalées sur milieu Luria Broth agar contenant 100 μ g/mL de X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) et 100 μ g/mL d'ampicilline. Les colonies de couleur bleue sont testées par PCR à l'aide des amorces EM194 et EM195 et envoyées à séquencer. Les multi-insertions sont vérifiées avec le triplet d'amorces EM111, EM112 et EM113 (Powell et al., 1994). La stratégie de clonage avec les amorces utilisées est résumée dans le tableau 4. Pour les fusions *ynfB* et *ompX*, une méthode alternative a été utilisée pour construire les fusions à partir de la souche NM580.

Pour remplacer le promoteur endogène de *mgtA* par le *PlacUV5* et pour le mutant de délétion de *mgtS* (anciennement *yneM*), une stratégie de mutagenèse dirigée a été utilisée avec le système λ_{red} (Datsenko et Wanner, 2000).

La polymérase T7, le dégradosome ARN, les protéines NusA et NusG ont été purifiés comme décrits précédemment (Rabhi et al., 2011). La polymérase d'ARN de *E. coli* a été achetée chez Epicentre et la bicyclomycine chez Santa Cruz Biotech.

2.2. SHAPE NATIF

Les réactions de SHAPE ont été réalisées comme décrit précédemment (Heppell et al., 2011). Les complexes natifs d'élongation de la transcription ont été obtenus par incubation des solutions suivantes : 20 mM Tris-HCl, pH 8,0 ; 100 mM KCl ; 20 mM MgCl₂ ; 0,5 μM de matrice d'ADN ; 1,6 μM de polymérase d'ARN de *E. coli* et 1,7 μM de facteur σ^{70} à 37°C pendant 3 min. Les réactions de transcription ont été complétées en ajoutant les rNTP à 1 μM à 37 °C pendant 10 min dans le même tampon contenant la concentration de MgCl₂ indiquée. Les ARN ont par la suite été incubés à 10% de DMSO ou 150 mM de NAI pour 15 min à 37°C. Les réactions de transcription ont été arrêtées par filtration sur colonne de Sephadex G50 et les ARN purifiés collectés dans du tampon TE 0,5X. Les transcrits modifiés ont été incubé avec une amorce d'ADN radiomarquée à 65 °C pour trois min, puis refroidis à 22°C pour 5 min. Les réactions de transcription inverse ont été réalisées comme décrit précédemment (Heppell et al., 2011) et les produits ont été séparés sur gel d'électrophorèse dénaturant 6% (19 :1).

2.3. CARTOGRAPHIE ENZYMATIQUE

Les transcriptions sont réalisées en deux étapes. La première se fait avec 20 mM de Tris-HCl pH 8,0 ; 20 mM MgCl₂ ; 20 mM NaCl ; 14 mM 2-mercaptoéthanol et 0,1 mM EDTA, le trinuécléotide UAC à 5 μM, les nucléotides ATP et GTP à 1,25 μM, 1 pmol de la matrice d'ADN contenant *PlacUV5*, la polymérase d'ARN de *E. coli* Epicentre et du [α -32P] UTP. Les complexes d'élongation sont alors purifiés sur colonne de Sephadex G50 et la

transcription est continuée dans le tampon contenant la concentration indiquée et les 4 rNTP. La RNase T1 est diluée au 1/1000^e et la RNase A au 1/10 000^e. La réaction est incubée 1 min à 37°C avant d'être arrêtée avec l'ajout de phénol/chloroforme/alcool isopropanyl (25:24:1). Les transcrits digérés sont enfin séparés sur gel d'électrophorèse dénaturant.

2.4. ESSAIS B-GALACTOSIDASE

Les essais cinétiques pour les expériences de β -galactosidase ont été réalisés comme décrit précédemment (Caron et al., 2012) et les fusions utilisées sont décrites au paragraphe 2.1.1. Brièvement, une culture bactérienne incubée 12 à 16 heures dans le milieu minimum M63 supplémenté de 0,2 % de glucose est diluée à une DO₆₀₀ de 0,02 dans 50 mL de milieu frais à la concentration de MgSO₄ indiquée. La culture est incubée à 37 °C pendant 9h, jusqu'à ce que la phase stationnaire soit atteinte. Les expériences avec le facteur Rho ont été réalisées dans 3 mL de culture (M63 Glucose, aux concentrations de Mg²⁺ indiquées). La bicyclomycine (BCM) a été ajoutée à 25 µg/mL tel qu'indiqué. Le TPP est ajouté à 500 µg/mL quand indiqué.

2.5. TRANSCRIPTIONS À CYCLE UNIQUE

Les transcriptions *in vitro* ont été réalisées dans un tampon composé de 20 mM de Tris-HCl pH 8,0 ; 20 mM MgCl₂ ; 20 mM NaCl ; 14 mM 2-mercaptoéthanol et 0,1 mM EDTA. La matrice d'ADN, le facteur σ^{70} et la polymérase d'ARN sont incubés à 37°C pour 5 min. Le tri-nucléotide UAC, nucléotides ATP/GTP et [α -32P] UTP sont ajoutés par la suite et la réaction est incubée à 37°C pour 8 min, donnant ainsi des complexes d'élongation bloqués à la position 26. L'échantillon est filtré sur colonne de Sephadex G50 pour retirer les nucléotides libres. Les réactions de transcriptions sont reprises ensuite dans le tampon précédent à la concentration de

MgCl₂ indiquée en ajoutant les quatre nucléotides avec de l'héparine pour ne permettre qu'un cycle de transcription.

2.6. IMMUNOBUVARDAGE DE TYPE NORTHERN

Les cultures bactériennes ont été incubées à 30 ou 37°C dans le milieu minimal M63 supplémenté avec 0,2% de glucose et à 20 µM de MgSO₄ jusqu'à la fin de la phase exponentielle (DO₆₀₀~0,8). Les cellules sont ensuite centrifugées et suspendues dans de l'eau stérile et les ARN totaux sont extraits immédiatement au phénol chaud (Caron et al., 2012). Le MgSO₄ est ajouté pour atteindre 50 µM de concentration finale. La bicyclomycine est ajoutée à 25 µg/mL au temps indiqué. Pour la détermination de la demi-vie, la transcription est bloquée par l'ajout de rifampicine à 250 µg/mL. Les expériences d'immunobuvardage de type ont été réalisés comme décrits précédemment (Caron et al., 2012). Les sondes ont été générées par PCR (Tableau 3 pour les amorces) puis transcription T7 radioactive. L'IPTG est ajouté à 1 mM quand indiqué.

2.7. DÉGRADATION *IN VITRO*

Les matrices d'ADN correspondant à la région 5' de *mgtA* ont été transcrites *in vitro* avec la polymérase T7 purifiée. Les transcrits ont été radiomarqués en 3' et utilisés pour les essais de dégradation *in vitro* décrits précédemment (Caron et al., 2012).

2.8. CHIP-QPCR

La souche $\Delta P_{mgtA}::PlacUV5$ a été incubée dans le milieu minimum M63 supplémenté de 0,2 % de glucose à 20 µM de MgSO₄ à 37 °C jusqu'à la fin de la phase exponentielle (DO₆₀₀ nm~

0,8). La co-immunoprécipitation de la chromatine avec la polymérase d'ARN a été réalisée comme décrit précédemment (Stringer et al., 2014) en utilisant un anticorps monoclonal de souris anti- β (NeoClone). Après la purification, des PCR en temps réel (qPCR) ont été réalisés avec l'instrument ABI 7500.

2.9. IDENTIFICATION DES SITES DE PAUSE

Les réactions de transcription ont été conduites en présence de 50 μ M de rNTP. Des aliquots ont été prélevés aux temps suivants : 30 sec, 1 min, 2 min, 3 min, 5 min 7 min, 10 min, 30 min et 60 min. Une réaction de *chase* est réalisée en ajoutant 1 mM rNTP. La détermination des demi-vies est faite comme décrit précédemment (Landick et al., 1996). Le séquençage transcriptionnel est réalisé comme décrit précédemment (Lemay et al., 2011).

2.10. PROFILAGE RIBOSOMIQUE

Les expériences de profilage ribosomique ont été réalisées comme décrites précédemment (Ingolia et al., 2013). La souche MG86 a été incubée jusqu'en fin de phase exponentielle à une concentration de 20 μ M de $MgSO_4$, la moitié de la culture a été additionnée de $MgSO_4$ pour atteindre la concentration finale de 50 μ M. Du chloramphénicol à 100 μ g/mL a été ajouté pour fixer les ribosomes sur l'ARN. 100 mL de culture ont été prélevés et les cellules ont été lysées à la presse de French (6 passages). Les échantillons ont été traités avec 3000 u/mL de micrococcale nucléase (MNase). Les polysomes sont ensuite séparés sur gradients de sucrose 10-50% préparés dans du tampon contenant 1 mM chloramphénicol et 2 mM DTT, pendant 2h30 à 35000 rpm (rotation par min) à 4°C. Après fractionnement, les échantillons contenant les monosomes ont été traités avec un mélange de phénol/chloroforme 5 :1 chauffé à 65°C. Après précipitation, les ARN ont été séparés sur gel d'acrylamide 15%, 19 :1. Les bandes comprises entre 20 et 50 nucléotides ont été découpées et placées dans un électro-élueur pour

extraire les ARN du gel. Les ARN ont été traités avec le kit RiboZéro afin d'enlever les ARN ribosomiques. Les étapes suivantes ont été effectuées comme décrites (Ingolia et al., 2013). Les libraires d'ADN complémentaires réalisées ont été envoyées à la plateforme de séquençage (MiSeq, « paired-end » 50 pb) du PRAC à l'Université de Sherbrooke. Les résultats ont été analysés comme indiqué au paragraphe 2.12. Le nombre de lectures obtenu pour chaque échantillon est de 8,874,790 pour l'échantillon faible Mg^{2+} n°1, 8,062,712 pour l'échantillon faible Mg^{2+} n°2, 9,181,342 pour l'échantillon fort Mg^{2+} n°1 et 7,650,828 pour l'échantillon fort Mg^{2+} n°2.

2.11 TRADUCTION *IN VITRO*

Les essais de traduction *in vitro* ont été réalisés comme décrit précédemment (Chauvier et al., 2017), en utilisant le kit PURExpress de New England Biolabs. Pour vérifier que l'ajout d'EDTA et de TPP n'affecte pas l'efficacité de transcription-traduction, une matrice d'ADN contenant les 934 premiers codons de *lacZ*, suivis d'un terminateur T7 et sous contrôle du *PlacUV5* a été obtenue par PCR. L'EDTA a été ajouté à 1 mM final et le TPP à 25 μ M. Les transcriptions ont été réalisées avec 0.5 U de PdARN commerciale (Epicentre). Les réactions sont arrêtées en plaçant les tubes sur glace pour 10 min puis par précipitation avec 4 volumes d'acétone. Les échantillons sont par la suite résolus sur gels SDS-PAGE.

2.12. ANALYSE DES DONNÉES DE SÉQUENÇAGE

Les données obtenues ont été analysées en utilisant le site Galaxy (<https://usegalaxy.org/>). La qualité des reads est vérifiée par l'outil « FastQC » et améliorée grâce aux outils « Trimmomatic » (Bolger et al., 2014) et « TrimGalore! » (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Ils sont par la suite alignés au génome de *E. coli* K12 MG1655 avec l'outil « Map with Bowtie for Illumina » (Langmead

et al., 2009). Le fichier SAM obtenu est converti en BAM avec « SAM-to-BAM » (Li et al., 2009) puis en fichier bedgraph visualisable sur le site <http://microbes.ucsc.edu/> avec l'outil « Create a Bedgraph of genome coverage » (Quinlan et Hall, 2010). Le fichier BAM est également utilisé pour compter le nombre de lecture par gène et intergène comme décrit précédemment (Lalaouna et al., 2017). Le tableau obtenu permet alors d'analyser les données comme décrit au chapitre 5, notamment grâce à la suite bio-informatique DESeq2 (Love et al., 2014) pour l'analyse de l'expression différentielle.

CHAPITRE 3

LE RIBORÉGULATEUR *MGTA* CHEZ *E. COLI*

3.1. LA RÉGION 5' NON TRADUITE DE *MGTA* PEUT FORMER DEUX STRUCTURES ALTERNATIVES

La première étude du riborégulateur *mgtA*, chez *S. enterica*, a établi la structure du riborégulateur d'après des prédictions *mfold* (Cromie et al., 2006). L'équipe a par la suite confirmé ces prédictions par cartographie enzymatique avec la RNase T1. En utilisant la structure du riborégulateur chez *S. enterica* et *mfold*, une partie de la structure a été prédite chez d'autres organismes dont *E. coli* (Figure 11).

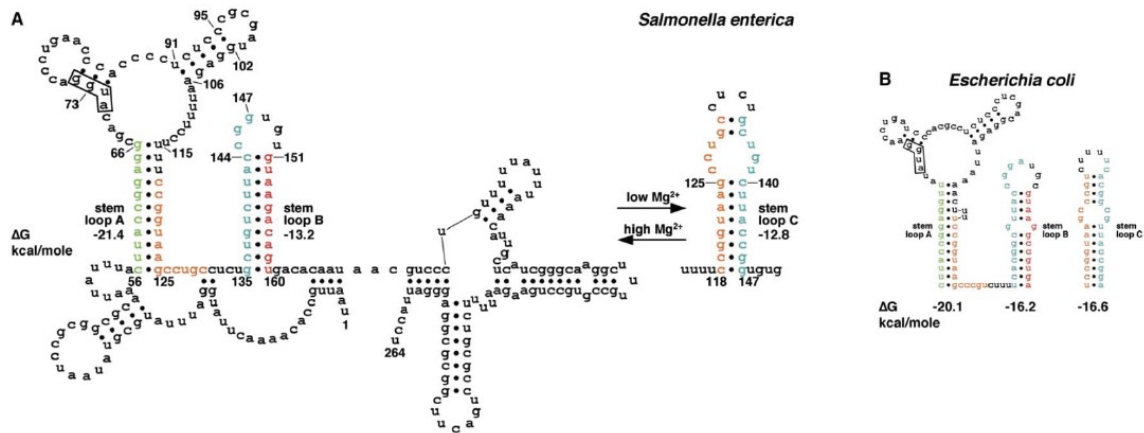


Figure 11. Prédictions des structures adoptées par le riborégulateur *mgtA* chez (A) *S. enterica* et (B) *E. coli*. Adapté de Cromie et al., 2006.

La région 5' non traduite de *mgtA* est longue de 262 nucléotides chez *E. coli*. La structure présentée sur la Figure 12A couvre la région de 54 à 163. Afin de déterminer la structure adoptée entre les nucléotides 164 et 262, un alignement de structures a été réalisé à l'aide des outils ClustalW et RNAalifold. L'alignement a été réalisé sur la région complète de *mgtA* provenant de 12 espèces proches avec ClustalW. L'alignement a ensuite été utilisé dans RNAalifold afin de trouver les structures conservées. L'alignement obtenu a permis de confirmer la structure dans la première région et d'établir une structure dans la deuxième région en 3' (Figure 12).

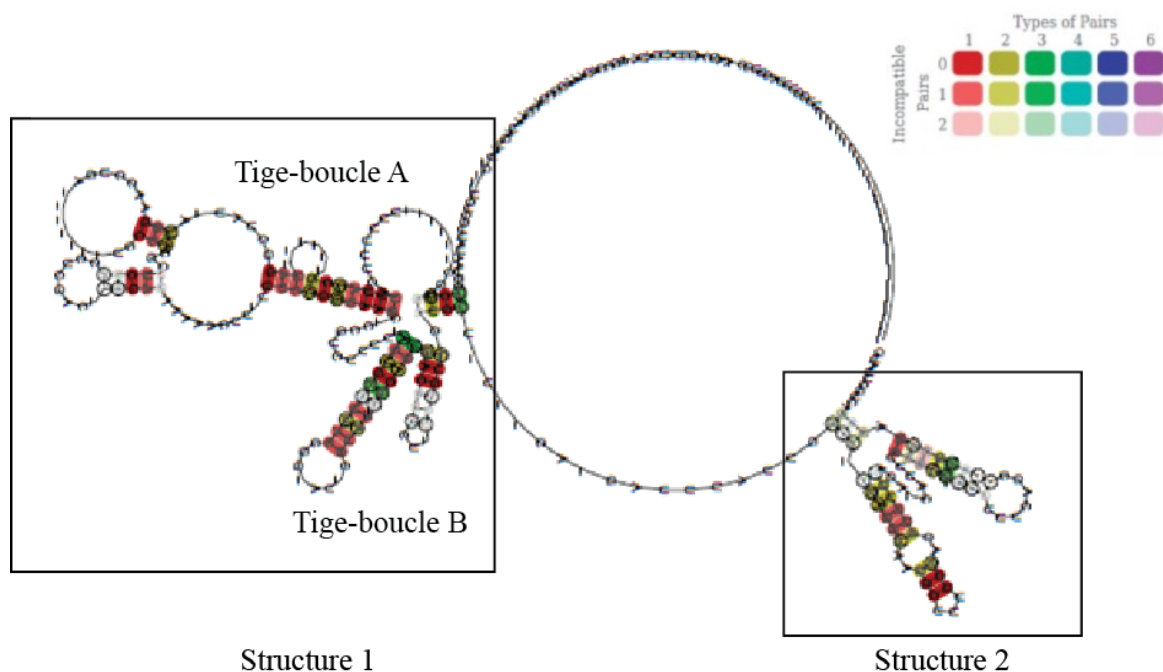


Figure 12. Alignement de structures de la région 5' de *mgtA*. Structures obtenues avec RNAalifold. Les tige-boucle A et B précédemment identifiées ont été retrouvées (structure 1) et une deuxième structure a été identifiée (structure 2). Les paires de base conservées sont entourées d'une couleur. La couleur indique le nombre de type de paires de base possible et la transparence indique la possibilité que cette paire de base ne se fasse pas.

Ces prédictions ont permis d'établir la structure inactive du riborégulateur présente en fort Mg^{2+} , par analogie avec les structures déterminées chez *S. enterica* (Cromie et al., 2006) (Figure 13). Néanmoins peu d'informations concernent la structure active présente en faible Mg^{2+} . La capacité de changer de structure suite à la liaison d'un ligand est une des caractéristiques majeures d'un riborégulateur. Il a donc été entrepris de confirmer les structures prédites et d'en apprendre d'avantage sur la structure active du riborégulateur. Dans un premier temps, pour confirmer la deuxième structure obtenue par alignement, une cartographie chimique a été réalisée. La technique de SHAPE (*Selective 2'Hydroxyl acylation Analyzed by Primer Extension*) a été choisie car elle ne nécessite pas de Mg^{2+} et donc ne devrait pas être affectée par la concentration en Mg^{2+} . Cette caractéristique est importante pour

pouvoir étudier la structure dans différentes concentrations de Mg^{2+} , ligand du riborégulateur *mgtA*. Comme plusieurs études ont montré l'importance du repliement cotranscriptionnel (Perdrizet II et al., 2012), la transcription des ARN pour la cartographie a été réalisée avec l'ARN polymérase d'*E. coli*. Le réactif de SHAPE (le NAI (2-methylnicotinic acid imidazolid)) a ainsi pu être ajouté directement après la transcription, permettant de cartographier la structure native du riborégulateur. Malheureusement, la transcription en faible Mg^{2+} est moins efficace dans les conditions utilisées, ne permettant pas de comparer les structures dans les deux concentrations de Mg^{2+} utilisées (Figure 13B). Les résultats obtenus permettent néanmoins d'étudier la structure du riborégulateur en fort Mg^{2+} . Les nucléotides les plus réactifs en bleu sont dans des régions non appariées de la deuxième structure : boucles 4 et 5, coude 3b et simples brins SB 4 et 5. D'autres nucléotides plus hauts dans le gel réagissent avec le NAI, mais la résolution ne permet pas de déterminer leur position. La technique de SHAPE a ainsi permis de confirmer l'existence de la deuxième structure *in vitro*. Afin d'observer des variations de structure avec le Mg^{2+} , une technique complémentaire a été utilisée, la cartographie enzymatique. Les conditions de transcription pour cette technique permettent d'avoir un rendement équivalent pour chaque condition de Mg^{2+} (Figure 13C). La concentration la plus basse de Mg^{2+} présente plus de bandes d'arrêt de transcription que les autres, mais l'ARN pleine longueur est aussi présent que dans les autres pistes. Deux RNases ont été utilisées pour une cartographie plus complète : la RNase T1 qui clive les G non appariés et la RNase A qui clive les C et les U non appariés. Les nucléotides auxquels l'ARN a été clivé sont indiqués respectivement en rouge et en vert. De manière générale les nucléotides clivés en fort Mg^{2+} confirment la première structure puisque ce sont des nucléotides présents dans les régions simple-brin et les boucles. Quelques changements sont observables avec le Mg^{2+} . Les C124 et C125 sont uniquement clivés à de plus fortes concentrations en Mg^{2+} . D'après les structures ces deux C se trouvent dans une région non appariée dans la structure inactive (fort Mg^{2+}) et dans la tige-boucle C dans la structure active (faible Mg^{2+}). Le U130 est clivé aussi bien en fort qu'en faible Mg^{2+} . Dans la structure inactive il se situe dans une région simple-brin et dans une boucle dans la structure active, ce qui explique sa disponibilité dans les deux conditions. Bien que les changements décrits ne

soient que des protections en faible Mg^{2+} , des protections en fort Mg^{2+} sont également observables dans le haut du gel, mais la résolution ne permet pas d'identifier les nucléotides impliqués. Afin de déterminer la structure de la région en amont de la tige-boucle C de la structure active, une prédiction avec *mfold* a été établie. Les deux structures les plus stables sont présentées à la Figure 14. Les résultats de cartographie permettent de déterminer laquelle des deux est la plus probable. En effet, les nucléotides C75 et C76 sont protégés en faible Mg^{2+} et ils sont appariés seulement dans la structure A. De plus le G86 est clivé indépendamment de la concentration en Mg^{2+} or il est apparié dans la structure B et non dans la structure A. La structure A semble ainsi la plus probable et c'est donc elle qui est représentée sur la Figure 13A. Pour la région en 3' de la tige-boucle C, la seule structure obtenue est identique à la deuxième structure en fort Mg^{2+} .

Les résultats de cartographie montrent ainsi que la région 5' non traduite peut adopter deux conformations en fonction de la concentration en Mg^{2+} lors de la transcription. Pour confirmer la présence d'un riborégulateur, il faut également que cette région régule l'expression du gène en aval selon la disponibilité en Mg^{2+} .

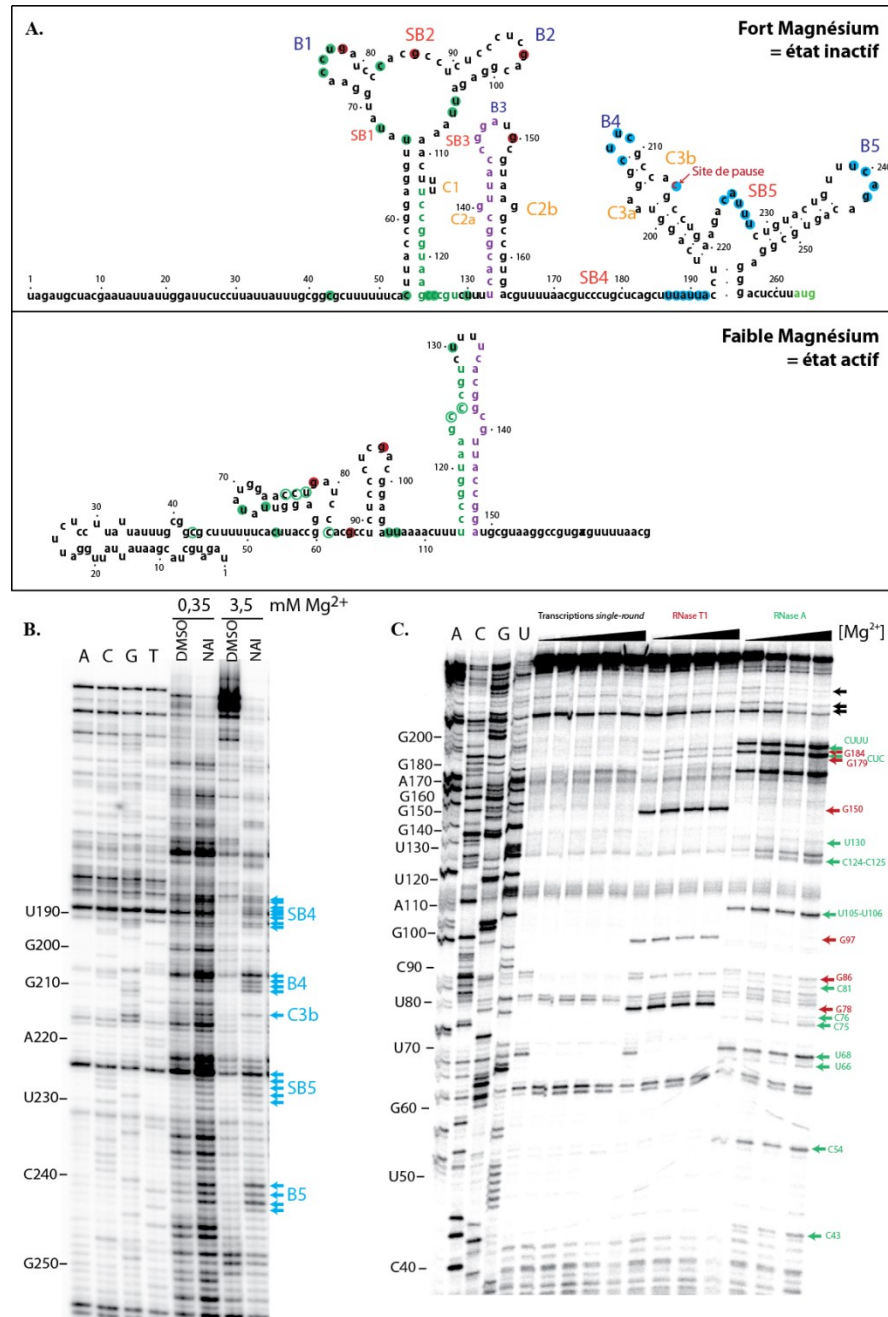


Figure 13. Cartographie de la structure du riborégulateur *mgtA* chez *E. coli*. A) Structures prédites. Les nucléotides qui réagissent en SHAPE (B) sont entourés d'un disque bleu. Les nucléotides clivés par la RNase A sont entourés d'un disque vert. Ceux clivés par la RNase T1 sont entourés d'un disque rouge. Les nucléotides entourés d'un cercle vert sont protégés en faible Mg²⁺. B : Boucle ; SB : Simple brin ; C : Coude. B) Cartographie par SHAPE. Une échelle de séquençage avec des didéoxyribonucléotides a été réalisée (4 pistes à gauche). Les régions simple-brin notées SB et les boucles notées B sont indiquées. C) Cartographie enzymatique par les RNases A et T1. Les échelles ont été produites par transcription avec des nucléotides 3'-O-méthyl. Les nucléotides clivés sont indiqués à droite. Les concentrations de Mg²⁺ utilisées pour la transcription sont 0,35-0,5-1-3,5 mM.

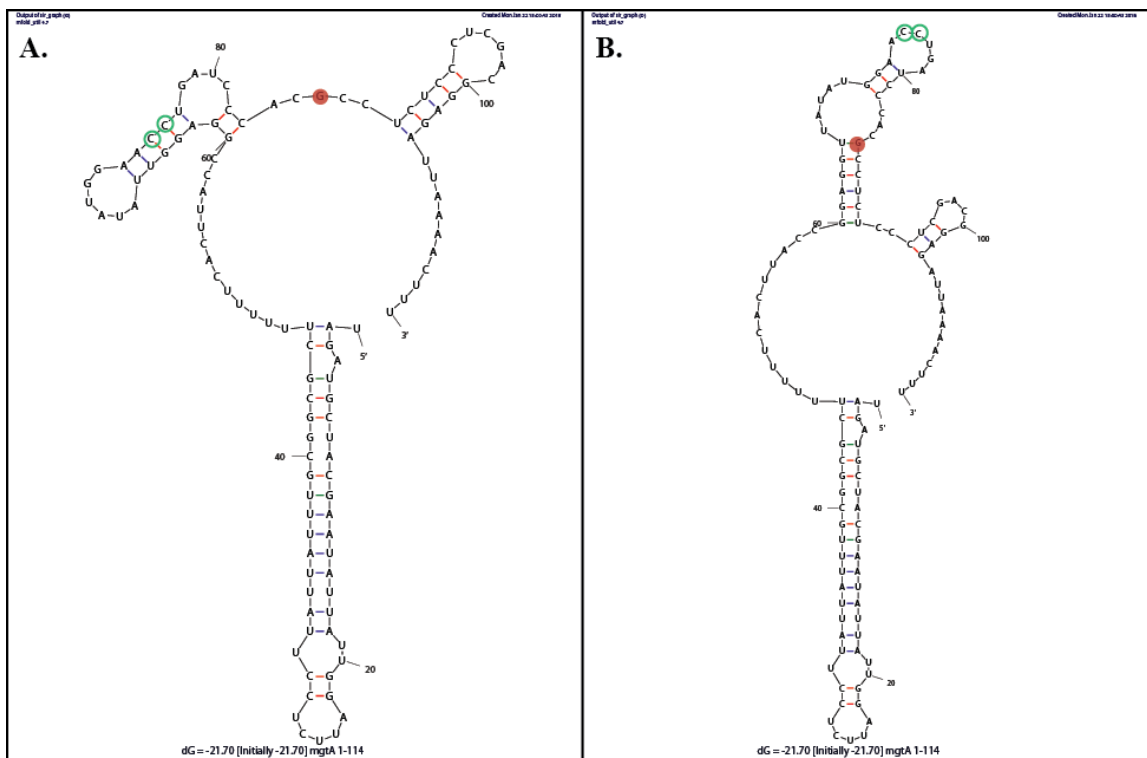


Figure 14. Deux prédictions mfold pour la structure active. Des résultats de cartographie (Figure 13) ont été reportés sur les structures pour déterminer la plus probable.

3.2. LA RÉGION 5' NON TRADUITE RÉGULE L'EXPRESSION DE *MGT*²⁺ AU NIVEAU DE L'ARN^m

Pour évaluer l'expression de *mgtA* *in vivo*, des fusions avec le gène rapporteur *lacZ* ont été construites (Figure 15A). Comme le promoteur *mgtA* est connu pour être régulé par le Mg^{2+} , une fusion transcriptionnelle contenant la région promotrice de *mgtA* et *lacZ* avec sa région d'initiation de la traduction a été construite pour avoir un contrôle positif de la régulation par le Mg^{2+} . Afin d'observer uniquement la régulation par le riborégulateur et pas par le promoteur, les fusions *mgtA* ont donc été construites avec un promoteur constitutif (*PlacUV5*). Pour vérifier que ce promoteur n'est pas dépendant du Mg^{2+} une fusion transcriptionnelle contenant uniquement celui-ci a également été réalisée. Enfin une fusion traductionnelle

MgtA'-LacZ a été construite pour suivre l'expression de *mgtA*. Cette fusion contient le promoteur *PlacUV5*, la région 5' non traduite de *mgtA* et les 8 premiers codons de *mgtA* fusionnés en phase avec le cadre de lecture de *lacZ*.

Par la suite, les conditions de Mg^{2+} et de croissances ont dû être déterminées. Pour faire varier la concentration intracellulaire en Mg^{2+} , la concentration en Mg^{2+} d'un milieu minimum classique (M63) a été modifiée. Les concentrations utilisées dans la littérature varient de 1 μM à 10 mM. Aucune croissance n'a été observée à 1 μM de Mg^{2+} , les concentrations testées varient donc de 10 μM à 10 mM, incluant 1 mM qui est la concentration du milieu M63 classique. Les courbes de croissances montrent de grandes variations avec le Mg^{2+} (Figure 15B). Afin de minimiser l'impact de la croissance, des concentrations de Mg^{2+} proches pour lesquelles une régulation peut être observée ont été choisies. Les essais β -galactosidase avec la fusion traductionnelle montrent que la croissance a également un effet sur la régulation : en phase exponentielle de croissance, il y a peu de différence d'activité entre les conditions de Mg^{2+} ; en fin de phase exponentielle, l'activité augmente drastiquement en faible Mg^{2+} alors qu'elle reste stable en fort Mg^{2+} (Figure 15C). Finalement, les conditions choisies pour les essais suivants sont : 20 μM pour la condition faible Mg^{2+} , 50 μM pour la condition faible Mg^{2+} et la phase de transition entre la phase exponentielle et la phase stationnaire. Dans ces conditions, la fusion promotrice de *mgtA* présente une régulation, mais pas la fusion promotrice *PlacUV5* ; Le promoteur est donc bien constitutif (Figure 15D). La fusion traductionnelle montre une régulation plus importante que la fusion promotrice, suggérant que le riborégulateur joue un rôle majeur dans la régulation de l'expression de *mgtA*. De plus, une fusion transcriptionnelle, pour laquelle la traduction de *lacZ* est indépendante de celle de *mgtA*, présente également une régulation. Ce résultat montre que la régulation s'effectue au moins au niveau de l'ARNm (transcription et/ou dégradation). Un immunobuvardage de type *northern* a été réalisé pour confirmer cette régulation (Figure 15E). Les ARN totaux ont été extraits d'une souche dans lequel le promoteur de *mgtA* a été remplacé par *PlacUV5*. Une

diminution rapide de l'ARNm *mgtA* est observée après ajout de Mg^{2+} , ce qui confirme les résultats des essais β -galactosidase.

La régulation par le Mg^{2+} étant établie *in vivo*, le riborégulateur *mgtA* est donc fonctionnel chez *E. coli*. La suite du chapitre vise à déterminer le mécanisme par lequel la régulation s'effectue.

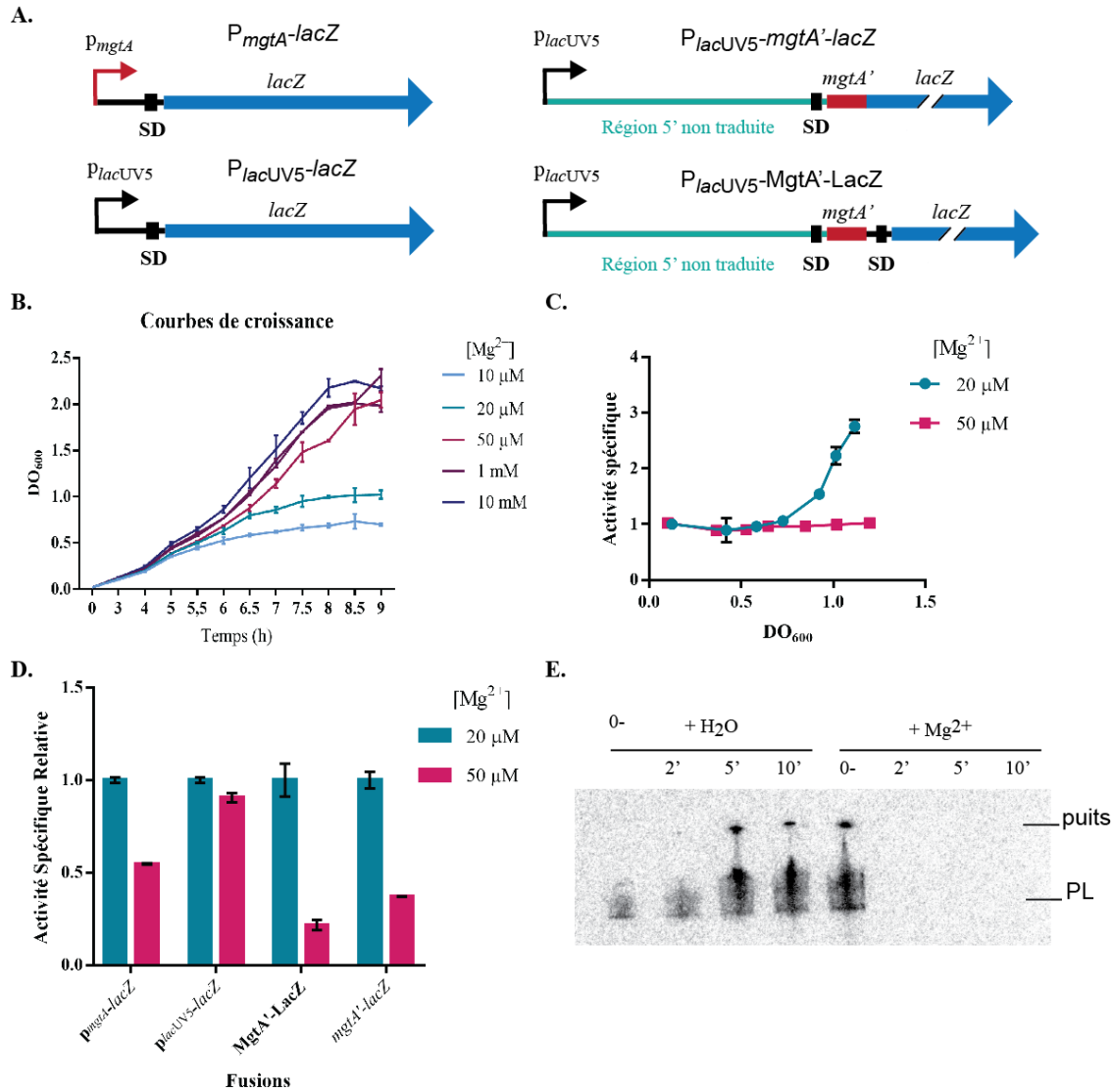


Figure 15. Régulation de l'expression de *mgtA*. A) Schéma des fusions *lacZ*: en haut à gauche, promoteur endogène *mgtA* ; en bas à gauche, promoteur *PlacUV5* ; en haut à droite, fusion traductionnelle *mgtA* avec ses 8 premiers codons, sous contrôle du *PlacUV5* ; en bas à droite, fusion transcriptionnelle *mgtA* avec ses 8 premiers codons, sous contrôle du *PlacUV5*. B) Courbes de croissance en M63 glucose, avec la concentration indiquée en Mg^{2+} . C) Activité β -galactosidase en fonction de la DO_{600} pour la fusion traductionnelle *mgtA* aux concentrations indiquées en Mg^{2+} . D) Activités obtenues en fin de phase exponentielle pour les fusions et concentrations de Mg^{2+} indiquées. Les résultats montrés sont les moyennes d'au moins deux expériences indépendantes avec les déviations standards. E) Immunobuvardage de type northern avec une sonde contre le riborégulateur *mgtA*. Les ARN ont été extraits en fin de phase exponentielle, à partir de la souche $\Delta P_{mgtA}::PlacUV5$ avant et après ajout d'eau ou de Mg^{2+} aux temps indiqués.

3.3. LE RIBORÉGULATEUR N'UTILISE PAS DE TERMINATEUR DE TRANSCRIPTION INTRINSÈQUE.

La régulation au niveau de l'ARNm peut s'effectuer au cours de la transcription par un arrêt de transcription prématuré ou post-transcriptionnellement par dégradation enzymatique. Comme il existe de nombreuses RNases chez *E. coli*, l'hypothèse d'un arrêt de transcription a été étudiée dans un premier temps. Il existe deux types de terminateurs : les terminateurs intrinsèques et les terminateurs dépendant de Rho. Ces premiers ne nécessitent pas de facteurs protéiques pour que la terminaison s'effectue. La terminaison peut ainsi s'effectuer lors d'une transcription *in vitro*. Des transcriptions à cycle unique ont ainsi été réalisées avec des concentrations croissantes en Mg^{2+} (Figure 16). La bande de terminaison la plus forte correspond à la séquence poly-U après la tige-boucle B. Cependant, la terminaison est plus forte en présence d'une faible concentration de Mg^{2+} , ce qui ne correspond pas aux résultats *in vivo*. Ce résultat pourrait être un artéfact dû à la faible concentration en Mg^{2+} qui pourrait faire décrocher plus facilement la polymérase d'ARN *in vitro*. Afin de vérifier cette hypothèse, un mutant du site poly-U pour un poly-A a été construit *in vivo*. Les essais β -galactosidase montrent que le mutant n'a aucun effet sur la régulation. Ces résultats suggèrent que la terminaison observée *in vitro* n'est pas responsable de la régulation observée *in vivo*.

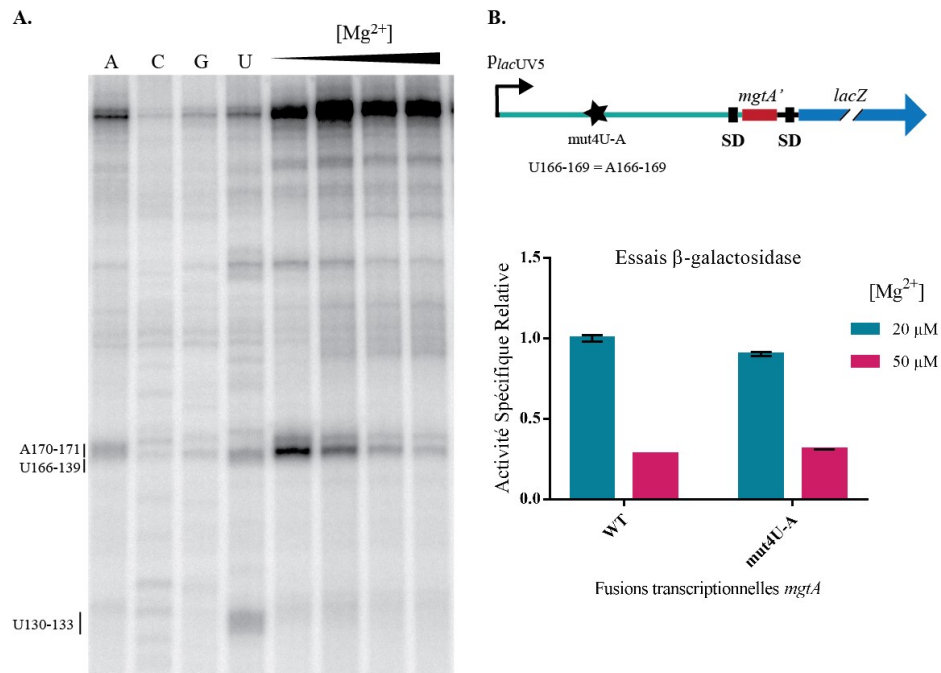


Figure 16. Le terminateur canonique n'est pas responsable de la régulation. A) Transcriptions à cycle unique aux concentrations de Mg^{2+} suivantes : 0,35 – 0,5 – 1 – 3,5 mM. Les échelles ont été réalisées avec des nucléotides 3'-O-méthyl. B) Essais β -galactosidase réalisés avec des fusions transcriptionnelles *mgtA* sauvage (WT) ou dans laquelle les U166-169 ont été mutés pour 4 A (*mut4U-A*). Les résultats montrés sont les moyennes d'au moins deux expériences indépendantes avec la déviation standard.

3.4. LE RIBORÉGULATEUR N'UTILISE PAS UN TERMINATEUR DÉPENDANT DE RHO

Les résultats du paragraphe précédent suggèrent que si le riborégulateur module la terminaison prématurée de la transcription pour réguler, il lui faut un facteur qui n'est pas présent dans les essais *in vitro*. Le seul facteur connu pour aider la PdARN à terminer est la protéine Rho. Afin d'évaluer l'importance de Rho dans la régulation un inhibiteur peut être utilisé : la bicyclomycine (BCM). Cette molécule empêche Rho d'avancer le long de l'ARN (Kohn et Widger, 2005). Pour observer l'effet de la BCM sur le niveau d'ARNm, un immunobuvardage de type *northern* a été réalisé avec la souche contenant la fusion transcriptionnelle et une sonde qui s'hybride dans *lacZ*. La BCM a été ajoutée 20 minutes avant l'ajout de Mg^{2+} et un contrôle dans lequel la BCM est remplacée par de l'eau a également été réalisé. Une

diminution de l'ARNm *mgtA-lacZ* est observée après ajout de Mg^{2+} aussi bien en présence de BCM qu'en son absence (Figure 17A). De plus, pour contrôler l'activité de la BCM une sonde contre l'ARNm *rho* a été hybridée. L'ARNm *rho* augmente 10 minutes après ajout de la BCM. Cet ARNm étant régulé par la protéine Rho, la BCM empêche sa terminaison prématurée. Il en résulte une augmentation de l'ARNm pleine longueur. Il semble donc que Rho ne soit pas impliqué dans la régulation de *mgtA* chez *E. coli*. Ce résultat étant à l'opposé de ce qui a été observé chez *S. enterica*, des essais β -galactosidase ont également été réalisés pour le confirmer. Les résultats obtenus ne montrent pas de différence en présence de BCM (Figure 17B). Une étude récente confirme également ces résultats avec une technique différente utilisant un mutant de Rho plutôt que la BCM (Bastet et al., 2017).

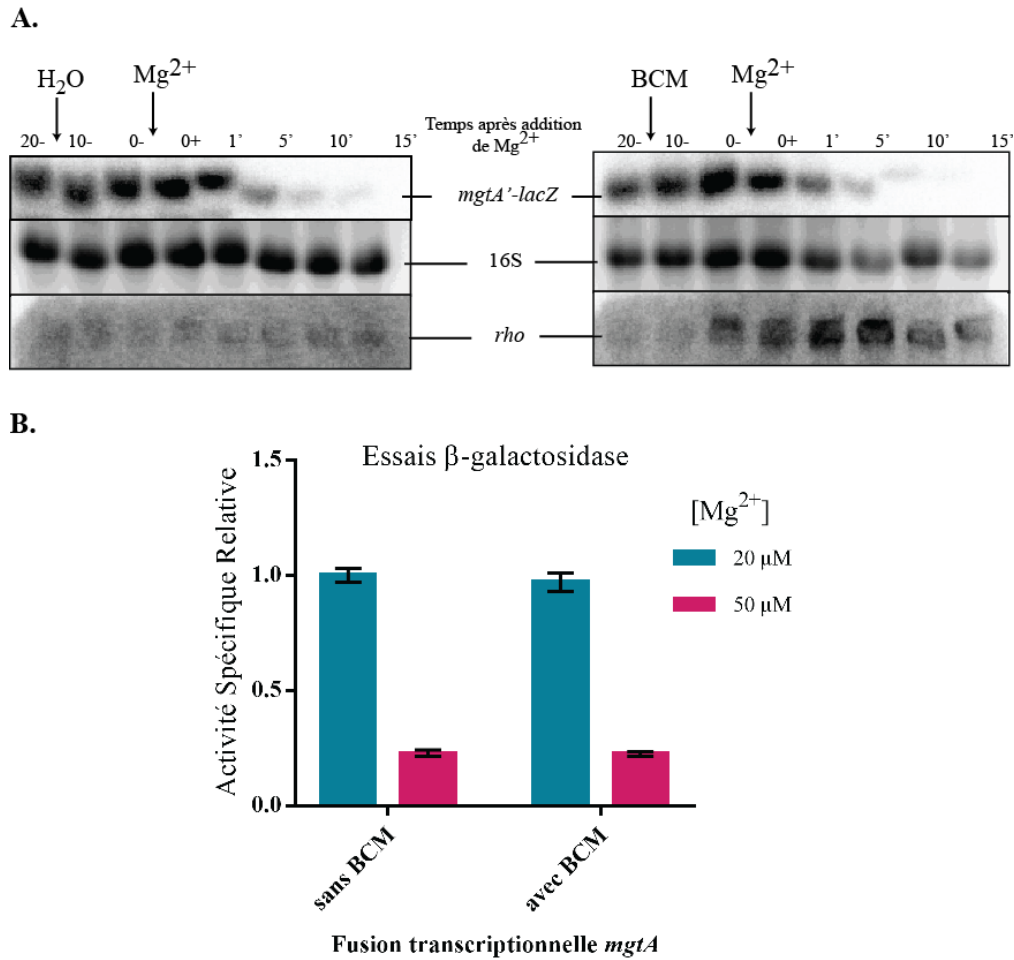


Figure 17. La protéine Rho n'est pas impliquée dans le mécanisme de régulation
 A) Immunobuvardage de type northern réalisé avec des ARN extraits de la souche contenant la fusion transcriptionnelle *mgtA*. Les sondes utilisées sont complémentaires à *lacZ*, les ARN 16S et *rho*. Les ARN ont été extraits aux temps indiqués par rapport à l'ajout de Mg²⁺ en présence d'eau ou de BCM. B) Essais β -galactosidase réalisés avec la fusion transcriptionnelle *mgtA* sauvage (WT) avec ou sans BCM aux concentrations en Mg²⁺ indiquées. La BCM a été ajoutée 30 min avant la récolte des cellules pour l'essai enzymatique.

3.5. LE DÉGRADOSOME N'EST PAS IMPLIQUÉ DANS LA RÉGULATION

Le riborégulateur ne semblant pas utiliser de terminateur de la transcription, l'hypothèse de l'implication d'une RNase responsable de la dégradation de l'ARNm en présence de fortes concentrations de Mg²⁺ a été envisagée. Il a de plus été remarqué que la demi-vie de l'ARNm

pleine longueur de *mgtA* est extrêmement faible (inférieure à 1 minute) (Figure 18). La rifampicine est ajoutée car elle bloque l'initiation de la transcription, ce qui permet d'observer la stabilité des ARN déjà transcrits (Campbell et al., 2001). La demi-vie des ARNm chez *E. coli* est d'environ 7 min en moyenne, bien au-dessus de celle de *mgtA* (Mohanty et Kushner, 1999; Selinger et al., 2003).

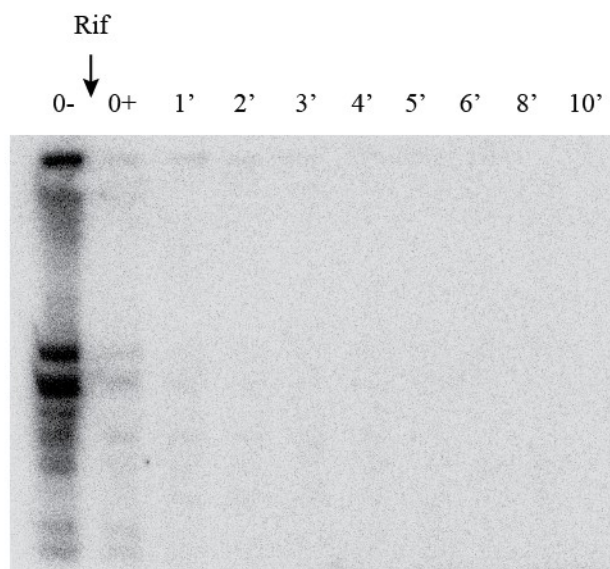


Figure 18. Demi-vie de l'ARNm *mgtA*. Immunobuvardage de type northern réalisé avec des ARN extraits de la souche sauvage, avant et aux temps indiqués après ajout de rifampicine (Rif). La sonde utilisée cible la région 5' de *mgtA*.

Chez *E. coli*, l'enzyme responsable de la dégradation de la majorité des ARNm est la RNase E en complexe avec d'autres protéines, appelé dégradosome ARN. Deux études ont montré que le dégradosome ARN pouvait être un acteur du mécanisme d'un riborégulateur, *lysC* chez *E. coli* (Caron et al., 2012) et *mgtA* chez *S. enterica* (Spinelli et al., 2008). La première étude a identifié un site de clivage de la RNase E : un site UCUU en amont d'une tige-boucle. Un site identique est retrouvé dans le riborégulateur *mgtA* chez *E. coli* aux positions 128 à 131 en amont de la tige-boucle B en fort Mg^{2+} . Dans un premier temps, le clivage du dégradosome au site UCUU a été vérifié *in vitro*. L'ARN transcrit et radio-marqué est incubé avec le

dégradosome purifié avant d'être mis à migrer sur gel de polyacrylamide. Une bande apparaît au niveau du site UCUU en présence du dégradosome (Figure 19A). Deux mutants ont été construits pour évaluer l'impact de la structure sur le clivage : un mutant ON pour lequel les nucléotides 54 à 61 ont été mutés pour leurs nucléotides complémentaires pour empêcher la formation de la tige-boucle A et ainsi favoriser la conformation active du riborégulateur ; et un mutant OFF dans lequel en plus de ces nucléotides, les nucléotides associés de 116 à 123 ont également été mutés pour leurs nucléotides complémentaires, stabilisant ainsi la tige-boucle A et donc la structure inactive. Comme attendu, le clivage est perdu dans le mutant ON, mais est encore présent dans le mutant OFF, suggérant que la structure contrôle le clivage par le dégradosome. Un mutant de délétion du site UCUU (Δ UCUU) a aussi été réalisé pour évaluer l'impact de la séquence. Le clivage disparaît dans ce mutant, confirmant l'importance du site UCUU. Pour confirmer ces résultats *in vivo*, la mutation Δ UCUU a été introduite dans la fusion transcriptionnelle. Étonnamment, les résultats des essais β -galactosidase montrent que la mutation n'affecte pas la régulation (Figure 19B). Néanmoins, l'activité est plus élevée dans ce mutant, suggérant que le site de clivage identifié *in vitro* est bien présent *in vivo*, mais ne fait pas partie du mécanisme de régulation employé par le riborégulateur. Pour confirmer ce résultat des mutants de la RNase E ont été utilisés *in vivo*. Dans un premier temps, le mutant catalytique thermosensible *rne-3071* (*rne-TS*), dans lequel la RNase E est inactive à 44 °C, a été utilisé (Apirion et Lassar, 1978). Un immunobuvardage de type *northern* montre que lorsque la RNase E est inactive l'ARNm *mgtA* est toujours dégradé alors que l'ARNr 5S n'est pas correctement mûré (Figure 19C). Dans un second temps, un mutant qui conserve l'activité ribonucléolytique de la RNase E tout en empêchant la formation du dégradosome ARN (Kido et al., 1996). Les résultats des essais β -galactosidase montrent que la régulation s'effectue toujours dans ce mutant (Figure 19D). La possibilité d'une double régulation à la fois par Rho et par la RNase E, l'un pouvant compenser lorsque l'autre est absent, a également été étudiée. Des essais β -galactosidase ont été réalisés avec la souche *rne-131* en présence de BCM pour inactiver en même temps Rho et la RNase E. Aucun effet n'a toutefois été observé sur la régulation, confirmant que ni Rho ni la RNase E ne sont impliqués dans la régulation chez *E. coli* (Figure 19D).

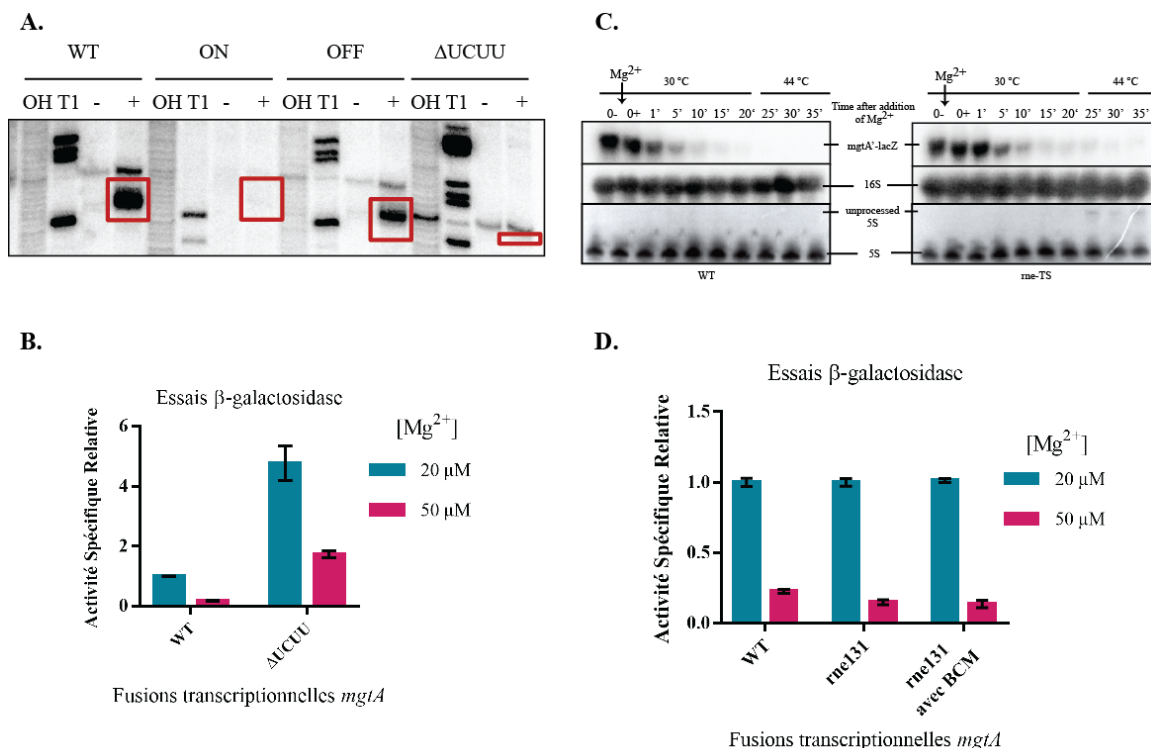


Figure 19. Implication du dégradosome dans la régulation. A) Essais de dégradation *in vitro*. Les transcrits T7 marqués radioactivement en 5', sauvage (WT) ou mutants, sont incubés avec le dégradosome purifié. OH : échelle réalisée avec du NaOH : clive tous les nucléotides ; T1 : échelle réalisée avec la RNase T1 : clive les G. B) Essais β-galactosidase réalisés avec la fusion transcriptionnelle *mgtA* sauvage (WT) ou mutante (ΔUCUU) aux concentrations en Mg²⁺ indiquées. C) Immunobuvardages de type northern réalisés avec la souche sauvage (WT) ou le mutant thermosensible de la RNase E (*rne-TS*) avec la fusion transcriptionnelle *mgtA*. Les cultures sont incubées à 30 °C puis sont mises à 44 °C pour inactiver la RNase E dans le mutant. Les sondes utilisées s'hybrident dans *lacZ*, les ARN 16S et les ARN 5S. D) Essais β-galactosidase réalisés avec la fusion transcriptionnelle *mgtA* dans les souches indiquées en présence ou en absence de BCM.

Le riborégulateur *mgtA* chez *E. coli* semble utiliser un mécanisme de régulation unique puisqu'il n'emploie aucun des trois mécanismes de régulations utilisés par les riborégulateurs (terminateur intrinsèque, terminateur dépendant de Rho, RNase E). *E. coli* possédant plus de 20 RNases et la dégradation d'un ARN résultant souvent de la combinaison de plusieurs RNases, l'identification de la ou les RNases impliquées dans la dégradation de l'ARNm *mgtA* sera un processus fastidieux (Arraiano et al., 2010). Avant d'entamer un tel projet la possibilité d'une régulation au niveau de la transcription sera ré-étudiée en premier lieu.

3.6. L'ÉLONGATION DE LA TRANSCRIPTION EST AFFECTÉE PAR LE Mg^{2+}

Afin d'étudier l'effet du Mg^{2+} sur la transcription, des expériences de ChIP-qPCR sur la polymérase d'ARN ont été réalisées. Des paires d'oligonucléotides (oligos) ont été conçues afin de couvrir la majorité du locus *mgtA*, du promoteur jusqu'à la séquence codante (Figure 20A). L'ADN a été extrait à partir de la souche $\Delta p_{mgtA}::PlacUV5$ cultivée en faible et fort Mg^{2+} . Les résultats obtenus pour chaque paire d'oligos et chaque condition de Mg^{2+} sont représentés sur la Figure 20B. Un effet notable du Mg^{2+} sur la présence de PdARN totale se traduit par une plus grande quantité de polymérase en faible qu'en fort Mg^{2+} (Figure 20C). Cet effet étant observé dès le promoteur (paire d'oligos n°1), l'effet du Mg^{2+} sur le promoteur a été vérifié en tirant sur le facteur σ^{70} (Figure 20D). Aucun effet sur son recrutement n'est observé confirmant que le promoteur *PlacUV5* ne répond pas au Mg^{2+} . Ce résultat suggère que la polymérase est recrutée de façon équivalente au promoteur, mais qu'elle décroche en fort Mg^{2+} peu après l'initiation. Ce résultat indique également qu'il est peu probable que l'augmentation de signal pour la PdARN soit due à un deuxième promoteur.

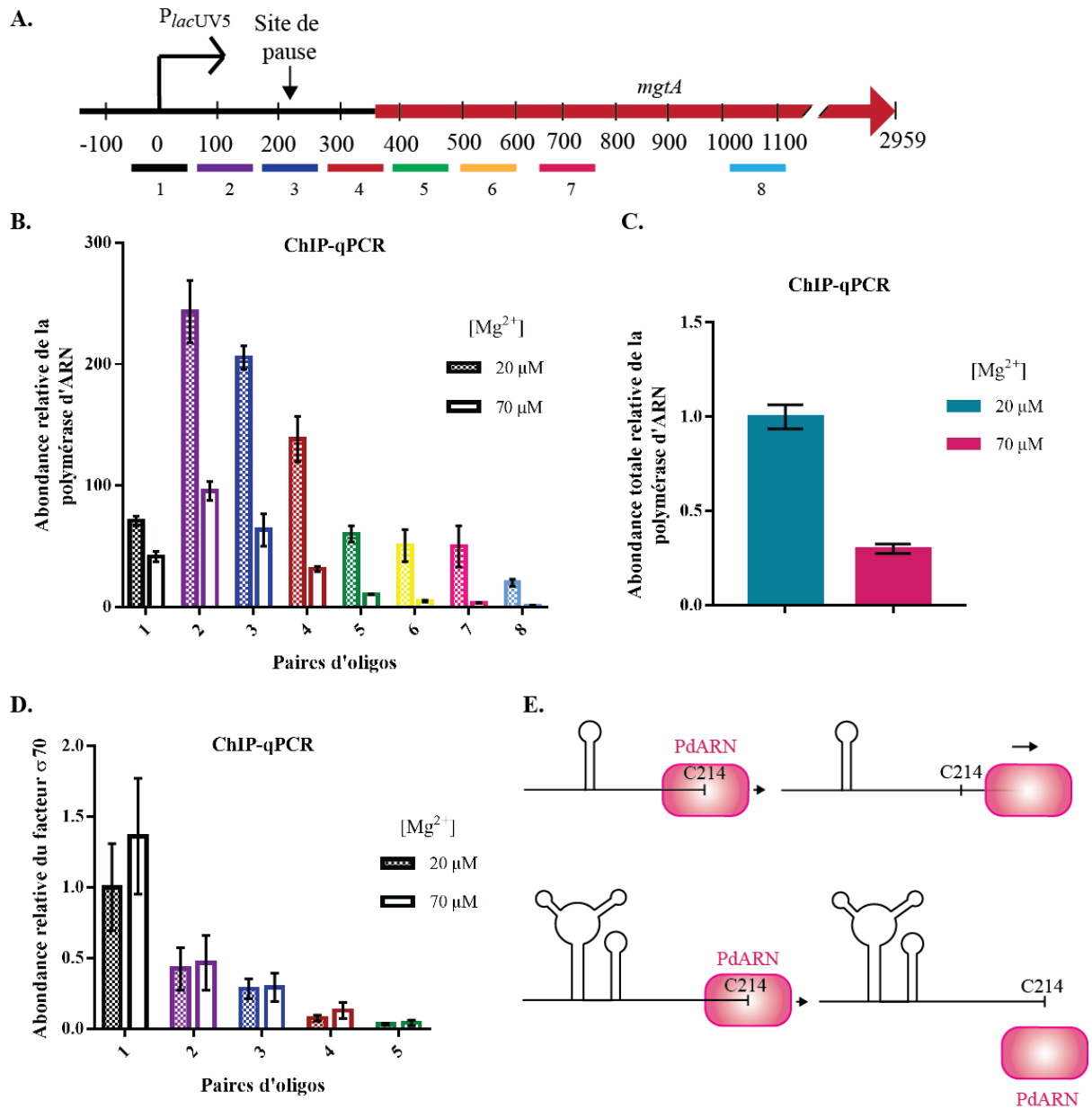


Figure 20. ChIP-qPCR de la polymérase d'ARN en faible et fort Mg^{2+} . A) Représentation schématique du locus *mgtA* et la position des paires d'amorces de qPCR. B) Résultats de ChIP-qPCR de la polymérase d'ARN relativisés avec un contrôle *rrnB*. Les couleurs correspondent aux paires d'amorces identifiées en A). Les résultats présentés sont la moyenne de trois expériences indépendantes avec leur écart-type. C) Total de l'abondance de polymérase présenté en B) pour chaque condition de Mg^{2+} relativisé à la condition faible Mg^{2+} . D) Résultats de ChIP-qPCR du facteur σ^{70} relativisés et normalisés avec la première paire d'amorces. Les couleurs correspondent aux paires d'amorces identifiées en A). E) Modèle proposé expliquant les résultats observés. En faible Mg^{2+} la polymérase d'ARN pause puis continue l'élongation de la transcription alors qu'en fort Mg^{2+} , elle décroche après avoir pausé. Les expériences de ChIP-qPCR ont été réalisées par le laboratoire du Pr Joseph Wade.

Étonnement, à la fois en faible et en fort Mg^{2+} , une augmentation de la présence de polymérase est observée pour la deuxième paire d'oligos, avant de diminuer pour les paires suivantes. Cette augmentation peut être expliquée par la présence d'un ou plusieurs sites de pause de la polymérase. Les sites de pause ont donc été déterminés *in vitro* en faible et fort Mg^{2+} (Figure 21A). Plusieurs sites ont été identifiés, dont un au C214 particulièrement fort et long, avec une demi-vie de 109 secondes en faible Mg^{2+} et de 151 en fort Mg^{2+} . Ces demi-vies, en accord avec celles trouvées pour la même pause chez *S. enterica*, sont beaucoup plus grandes que celle d'une pause classique comme celle de l'opéron *his*, de 8-9 secondes à la même concentration en rNTP (Hollands et al., 2014). La longueur et la dépendance au Mg^{2+} suggèrent un rôle potentiel de la pause dans la régulation de l'expression de *mgtA*, la terminaison de la transcription étant commune aux sites de pause (Weixlbaumer et al., 2013). Des mutants de la pause ont donc été construits pour observer l'effet de l'absence de pause. Les pauses pour ces mutants ont été déterminées et les demi-vies calculées sont de 98 secondes pour G205C, 78 secondes pour C214G, et 172 secondes pour G215A en fort Mg^{2+} (Figure 21B). Le mutant C214G a une pause d'une demi-vie six fois inférieure au WT et semble même perdre complètement la pause. Cette mutation a ensuite été introduite dans une fusion transcriptionnelle *in vivo*. Les résultats β -galactosidase ne montrent cependant aucun effet de l'absence de cette pause sur la régulation (Figure 21C). Ce résultat inattendu peut suggérer que malgré la disparition de la pause *in vitro*, celle-ci pourrait toujours exister *in vivo*, celle-ci pouvant être modulée par des facteurs de transcription tels que NusA et NusG. L'effet de ces deux protéines sur la pause a été observé *in vitro* : NusA augmente la pause alors que NusG la fait disparaître, comme décrit précédemment (Artsimovitch et Landick, 2000) (Figure 21D). L'absence d'effet du mutant C214G peut également s'expliquer par le fait que la mutation favorise la terminaison ou la dégradation du transcrit, indépendamment de la pause, ou que la pause n'a aucun rôle dans la régulation. Si la pause à C214 est responsable du profil de ChIP observé, la polymérase devrait être majoritairement présente dans la région amplifiée par la paire d'amorce 3 et non 2. Ce décalage peut s'expliquer par le fait que la polymérase pause mais décroche également à la position 214, décrochage qui se traduit par un

signal de qPCR plus faible. De plus, avec la polymérase à 214, il pourrait y avoir une accumulation de polymérases en amont, amplifiée avec la pause aux U112-114.

Bien que la technique de ChIP-qPCR ne permette pas de déterminer précisément le ou les sites de terminaison de la transcription, les résultats obtenus, complétés par la détermination des sites de pause, suggèrent un mécanisme de terminaison de la transcription au niveau de la deuxième structure du riborégulateur (Figure 20E). Étonnamment, cette terminaison n'est pas dépendante de Rho et n'est pas observée *in vitro*. Ces constatations suggèrent un mécanisme plus complexe qu'un riborégulateur classique. La présence d'un peptide leader observée chez *S. enterica* pourrait expliquer la complexité du mécanisme (Park et al., 2010) et son rôle chez *E. coli* sera donc étudié afin de compléter les connaissances sur le riborégulateur *mgtA* chez *E. coli*.

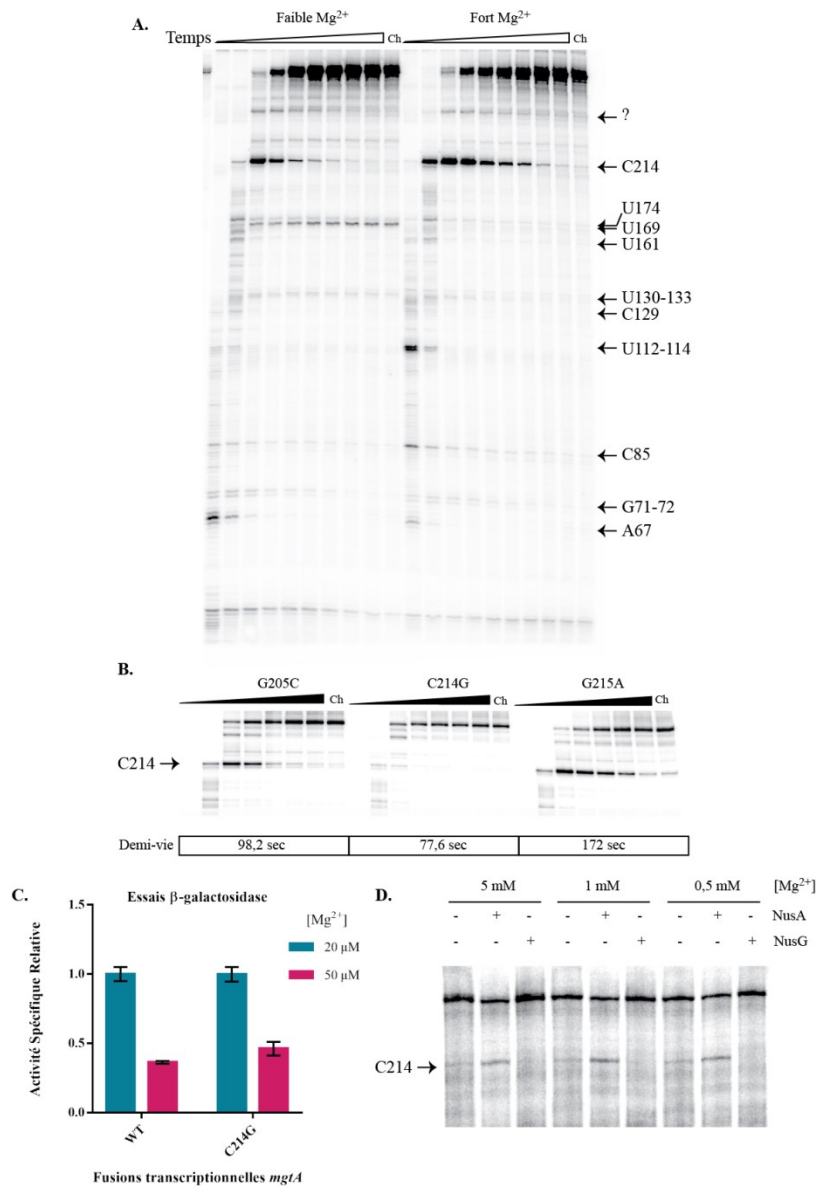


Figure 21. Les sites de pause de la polymérase d'ARN dans le riborégulateur *mgtA*. A) Transcriptions à cycle unique dans le temps en faible et fort Mg²⁺. Les ARN sont transcrits avec la polymérase de *E. coli*. Une première étape de transcription permet de marquer radioactivement le transcrit et changer le tampon de transcription pour un contenant soit 0,35 mM soit 3,5 mM. La concentration de rNTP est à 50 μM pour ralentir la transcription. La piste Ch correspond au « chase » : après le dernier prélèvement des rNTP sont ajoutés à 1 mM pour compléter tous les transcrits. Les sites de pauses sont cartographiés grâce à des échelles 3'O-méthyl non présentées. B) Transcriptions dans le temps des mutants mentionnés en fort Mg²⁺. Les temps de demi-vie sont présentés pour chaque mutant. C) Essais β-galactosidase réalisés avec la fusion transcriptionnelle *mgtA* sauvage (WT) ou mutante (C214G). Les résultats présentés sont la moyenne et écart-type d'au moins deux expériences indépendantes. D) Transcriptions à cycle unique réalisées en présence des facteurs mentionnés. La pause précédemment identifiée est indiquée.

CHAPITRE 4

LE PEPTIDE LEADER *MgtL*

4.1. PRÉSENCE D'UN CADRE DE LECTURE DANS LE RIBORÉGULATEUR

Une des particularités du riborégulateur *mgtA* est la présence d'un cadre de lecture localisé dans la région capable de changer de conformation. Plusieurs études de ce peptide leader ont été réalisées chez *S. enterica* mais leurs conclusions divergent et ne s'accordent pas pour savoir si le peptide leader fait partie du mécanisme du riborégulateur ou si c'est un deuxième mécanisme de régulation indépendant. Le peptide leader *mgtL* est conservé chez *E. coli* et son rôle dans la régulation de l'expression de *mgtA* sera étudié dans ce chapitre. Le cadre de lecture est identifié sur la Figure 22A pour les deux conformations. Le SD est dans les deux cas séquestré, mais semble moins accessible en fort Mg^{2+} . La séquence du peptide est de 17 résidus et contient 4 prolines et 4 arginines, ce qui représente une grande proportion. La séquence peptidique et la traduction du peptide pourraient donc avoir un rôle de régulation de l'expression de *mgtA*. Dans un premier temps, une fusion traductionnelle a été construite pour déterminer si le peptide pouvait être traduit *in vivo* chez *E. coli*. Les essais β -galactosidase montrent que cette fusion a une activité supérieure à celle de *mgtA*, suggérant que le peptide peut effectivement être traduit (Figure 22B). Bien que les essais aient été réalisés dans les deux conditions de Mg^{2+} , il est impossible de conclure quant à la régulation de l'expression de *mgtL* par le Mg^{2+} . En effet, la fusion *MgtL*'-LacZ n'inclue qu'une partie du riborégulateur qui n'est pas suffisante pour changer de conformation.

Le peptide pouvant être traduit *in vivo*, la possibilité qu'il fasse partie du mécanisme de régulation du riborégulateur a été envisagée. Dans ce cas, son expression pourrait être régulée par le Mg^{2+} . Les études précédentes ont conclu que l'expression de *mgtL* était activée par le

Mg^{2+} , à l'inverse de *mgtA*, mais elles ne se basent que sur les résultats obtenus avec une fusion traductionnelle de *mgtL*, qui ne contient pas le riborégulateur au complet. Pour ces raisons, une méthode a dû être développée pour répondre à cette première question.

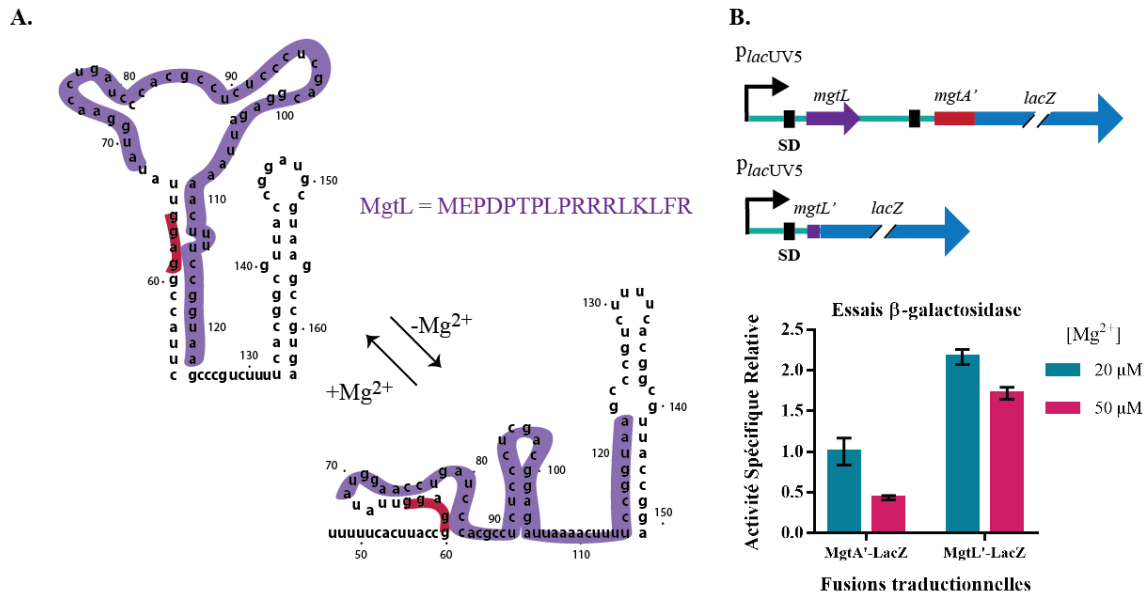


Figure 22. Une région du riborégulateur peut être traduite *in vivo*. A) Représentation du cadre de lecture en violet sur la structure du riborégulateur. La séquence peptidique de MgtL est également indiquée. La séquence SD est surlignée en rouge. B) Essais β-galactosidase réalisés avec des fusions traductionnelles de *mgtA* et *mgtL*. La fusion MgtL'-LacZ ne contient que le premier codon. Les résultats sont la moyenne et l'écart-type d'au moins deux expériences indépendantes.

4.2. LE PEPTIDE EST TRADUIT ET SON EXPRESSION EST RÉGULÉE PAR LE Mg^{2+}

Pour observer la régulation de l'expression de *mgtL* dans un contexte endogène, la technique de profilage ribosomique ou *ribosome profiling* a été choisie. Comme la purification des polysomes ne permettrait pas de distinguer entre un ARNm traduit uniquement au niveau de *mgtA* ou *mgtL* d'un ARNm traduit aux deux *loci*, une digestion des ARN non protégés par des ribosomes suivie par un séquençage haut débit est nécessaire. Le profilage ribosomique a été réalisé en faible et fort Mg^{2+} en fin de phase exponentielle avec la souche contenant le

promoteur *PlacUV5* à la place du promoteur endogène de *mgtA*. La comparaison de nombreux gènes aux profils distincts permettra d'évaluer la qualité des résultats obtenus. Des gènes non régulés par le Mg^{2+} tels que *rpoA*, codant pour la sous unité α de la polymérase d'ARN, *dnaE*, codant pour une sous unité de la polymérase d'ADN et *glmS*, un gène régulé par un riborégulateur, ne présentent pas de différence significative d'expression entre les deux concentrations de Mg^{2+} (Figure 23A, panneau de droite ; Figure 23B, gènes de maintenance). À l'inverse, des gènes régulés par PhoP, dont *phoP* lui-même, *mgrB* et *mgtS* sont significativement réprimés par le Mg^{2+} comme attendu (Figure 23B, régulon PhoP). Le gène *mgtA* sous contrôle du promoteur *PlacUV5* est également régulé par le Mg^{2+} , en accord avec les résultats obtenus avec des fusions au chapitre 3 (Figure 23A, panneau de gauche). Pour *mgtL*, la présence de nombreuses lectures le long de la région codante confirme que le peptide est traduit *in vivo*, en accord avec la fusion traductionnelle (Figure 22B). De plus, les lectures sont beaucoup moins nombreux (2,5 fois moins) en fort Mg^{2+} , suggérant que l'expression de *mgtL* est réprimée par le Mg^{2+} chez *E. coli*. Ce résultat était inattendu puisqu'il est en désaccord avec les études chez *S. enterica*. Cependant, c'est la première fois que la régulation de *mgtL* par le Mg^{2+} est observée directement, sans fusion ni étiquette. Le fait que l'expression du peptide leader est régulée par le même facteur que le gène en aval indique que la traduction de *mgtL* pourrait être centrale dans la régulation par le riborégulateur.

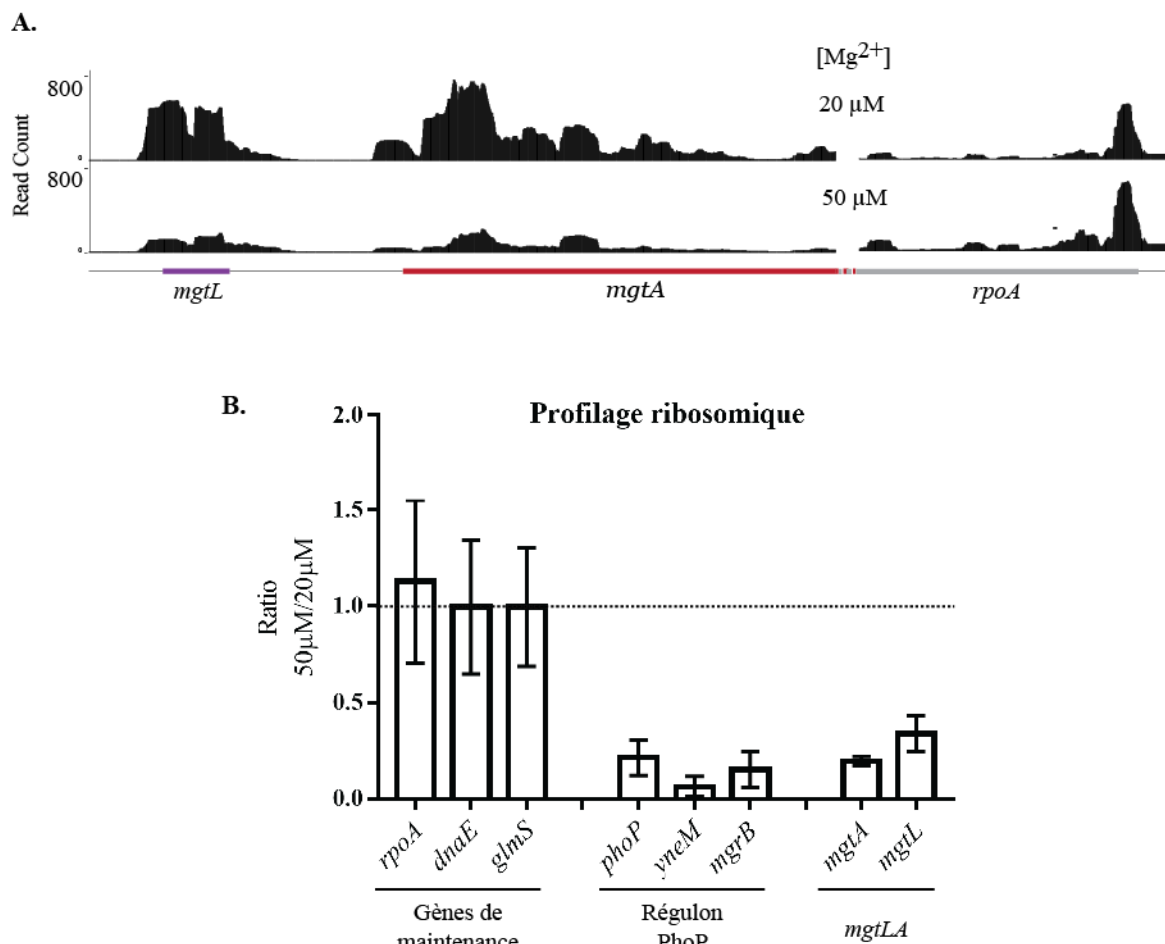


Figure 23. La technique de profilage ribosomique révèle l'expression et la régulation de *mgtL*. A) Résultats du profilage ribosomique obtenus avec la souche $\Delta P_{mgtA}::PlacUV5$ pour les gènes *mgtL*, *mgtA* et *rpoA*. Pour chaque position, le nombre de lecture est donné en ordonné en faible et fort Mg^{2+} . B) Ratios fort sur faible Mg^{2+} du nombre de lectures pour chaque gène identifié. Le groupe auquel appartiennent chaque gène est indiqué.

4.3. LA TRADUCTION DU PEPTIDE MGT_L EST NÉCESSAIRE À L'EXPRESSION DE *MGT_A*

L'expression de *mgtL* étant régulée dans le même sens que *mgtA*, il est possible que la traduction du peptide soit directement liée à l'expression de *mgtA*. Afin de tester cette hypothèse, des fusions *mgtA'-lacZ* dans lesquelles la traduction de *mgtL* est impossible ont été construites (Figure 24A). La première construction, pour laquelle le SD de *mgtL* a été muté

(mutSD), présente une perte de l'expression de *mgtA* et de la régulation, suggérant que la traduction de *mgtL* est nécessaire à la fois pour l'expression et la régulation de *mgtA* (Figure 24B). La deuxième construction, pour laquelle le codon d'initiation AUG a été muté pour un codon non-initiateur AGG, montre les mêmes résultats que mutSD, confirmant les premières observations (Figure 24B). Afin de vérifier que l'effet était bien dû à la perte de traduction de *mgtL*, l'AUG a également été remplacé par un GUG, autre codon d'initiation. Cette mutation a pour effet la perte d'expression, mais pas de la régulation (Figure 24B). Il a été envisagé que la perte d'expression soit due à une plus faible traduction de *mgtL* avec un GUG, ce qui a été confirmé par une fusion traductionnelle de *mgtL* avec un GUG, qui présente moins d'expression qu'avec l'AUG endogène (Figure 14C).

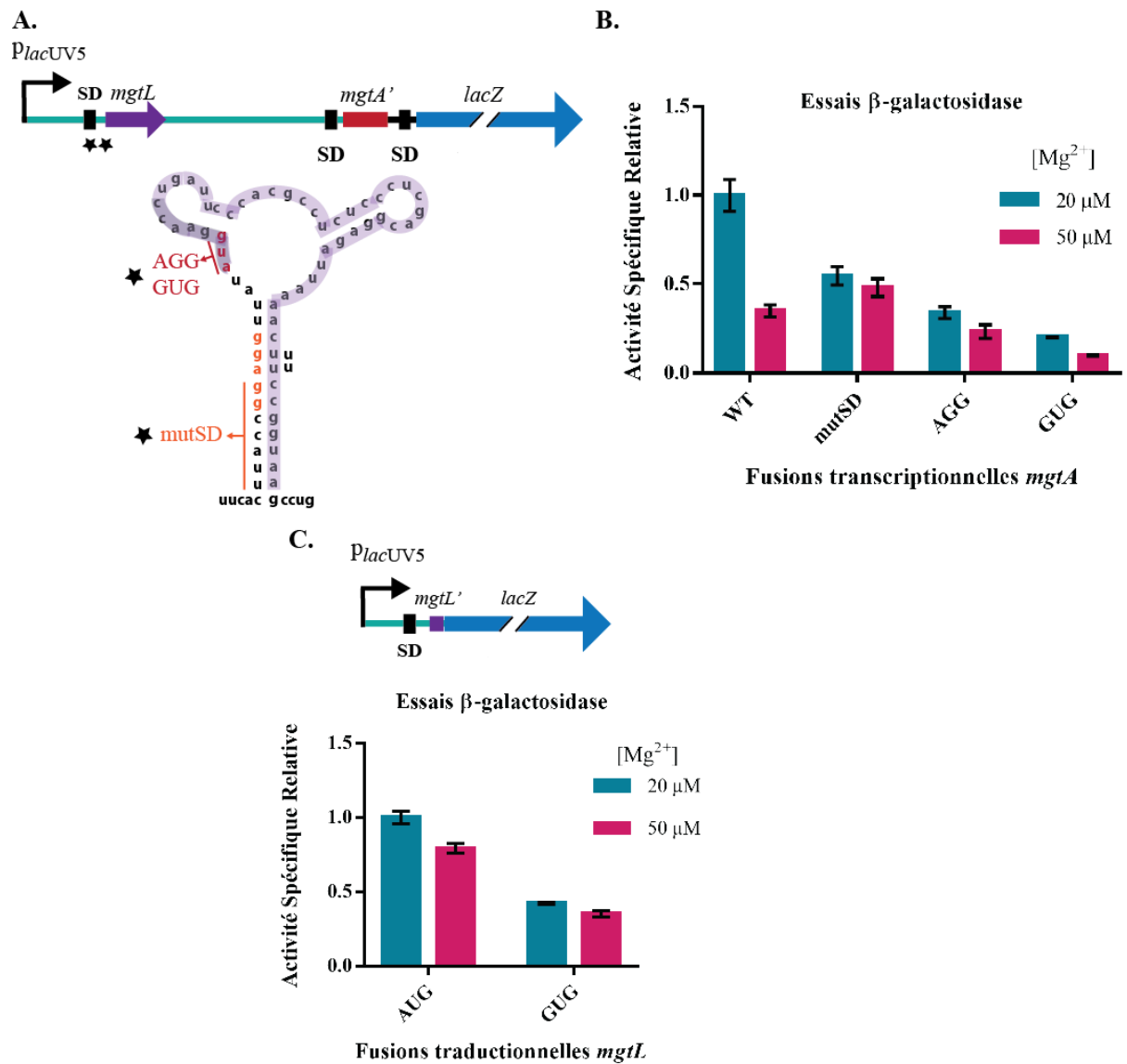


Figure 24. La traduction du peptide est nécessaire à l'expression et la régulation de *mgtA*. A) Schéma des fusions transcriptionnelles, les étoiles représentent les mutations identifiées sur la structure du riborégulateur. B) Essais β -galactosidase réalisés avec les fusions transcriptionnelles *mgtA* indiquées en faible et fort Mg²⁺. C) Essais β -galactosidase réalisés avec des fusions traductionnelles *mgtL* avec le codon sauvage AUG ou le mutant GUG. Les résultats sont la moyenne et l'écart-type d'au moins deux expériences indépendantes.

La traduction du peptide semble donc nécessaire à l'expression et à la régulation de *mgtA*. L'étape suivante vise à déterminer si c'est le processus de traduction de *mgtL* (et donc la

présence de ribosomes sur l'ARNm) ou si c'est le peptide MgtL produit qui a un effet sur l'ARNm *mgtLA*. Pour tester l'effet de MgtL, un plasmide de surexpression a été construit. Celui-ci contient la région 5' de *mgtA*. Pour évaluer son effet sur le niveau d'ARNm de *mgtA*, ce plasmide a été inséré dans une souche contenant une fusion transcriptionnelle *mgtA*. Dans cette fusion, l'AUG de *mgtL* est muté pour AGG (Figure 25A). La surexpression du peptide, vérifiée par immunobuvardage de type *northern* (Figure 25B), ne réprime pas l'expression de *mgtA* par rapport à une souche contenant le plasmide vide (Figure 25C). Ces résultats suggèrent que *mgtL* doit être traduit en *cis* et non en *trans* pour réguler l'expression de *mgtA*. La présence de ribosome dans cette région semble donc importante.

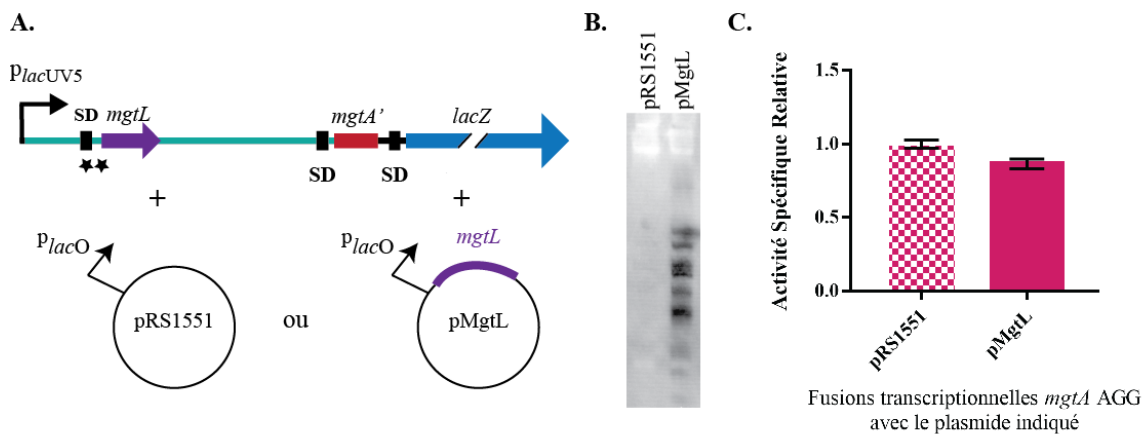


Figure 25. La surexpression de *mgtL* en trans n'affecte pas l'expression de *mgtA*. A) Schéma de la fusion transcriptionnelle *mgtA* avec le codon d'initiation de *mgtL* muté pour AGG et des plasmides transformés. pRS1551 est le plasmide vide et pMgtL le plasmide de surexpression de *mgtL*. B) Immunobuvardage de type northern d'ARN extraits à partir des souches contenant le plasmide indiqué en présence d'IPTG. La sonde utilisée cible le riborégulateur *mgtA* contenant *mgtL*. C) Essais β -galactosidase réalisés avec les souches contenant les plasmides indiqués. Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

4.4. LA RÉGULATION DE LA TRADUCTION DU PEPTIDE EST SUFFISANTE POUR RÉGULER L'EXPRESSION DE *MGT A*

Pour évaluer le rôle de la traduction de *mgtL* sur l'expression de *mgtA*, un outil a dû être développé afin de contrôler l'initiation de la traduction du peptide indépendamment du Mg^{2+} . Les riborégulateurs chez *E. coli* sont connus pour réguler l'expression du gène en aval en modulant l'accessibilité du SD et donc la traduction, faisant d'eux des candidats idéals pour contrôler artificiellement la traduction de *mgtL* (Barrick et Breaker, 2007). Cependant, des études récentes ont montré que plusieurs de ces riborégulateurs pouvaient réguler également au niveau de l'ARNm (Bastet et al., 2017). Afin de pouvoir étudier uniquement le rôle de la traduction de *mgtL* sur l'ARN *mgtA*, un riborégulateur contrôlant uniquement la traduction et non le niveau d'ARNm est nécessaire. Le riborégulateur *thiM* apparaît comme un bon candidat puisque des résultats obtenus au laboratoire ont montré que la séquence 5' non traduite de *thiM* seule ne contrôlait que la traduction (Bastet et al., 2017). Ce riborégulateur réprimant la traduction du gène en aval en présence de thiamine pyrophosphate (TPP), l'effet du TPP sur l'expression de *mgtA* a été vérifié dans un premier temps. Les essais β -galactosidase réalisés avec la fusion transcriptionnelle *mgtA* en présence et en absence de TPP montrent que l'expression de *mgtA* ne dépend pas du TPP (Figure 26B). Une fusion hybride contenant le riborégulateur *thiM* directement en amont de *mgtL* et du reste du riborégulateur *mgtA* et ses 8 premiers codons, comme pour la fusion *mgtA* sauvage, a été construite (Figure 26A). Dans cette fusion, l'initiation de la traduction de *mgtL* est directement sous le contrôle du TPP, ce qui permet d'évaluer si le niveau d'ARNm de *mgtA* peut être contrôlé par le TPP au lieu du Mg^{2+} . Les essais β -galactosidase réalisés avec cette fusion en présence et en absence de TPP montrent que le niveau d'ARN est moins important en présence de TPP (Figure 26B). Cette répression n'est pas observée pour la même fusion dans laquelle l'AUG de *mgtL* est muté (Figure 26B), ni pour une fusion transcriptionnelle ne contenant que le riborégulateur *thiM* (Figure 26B). Ce résultat indique qu'en contrôlant la traduction de *mgtL*, le niveau d'ARN de

mgtA est également régulé, ce qui suggère que la traduction de *mgtL* est responsable de la régulation de l'expression de *mgtA*.

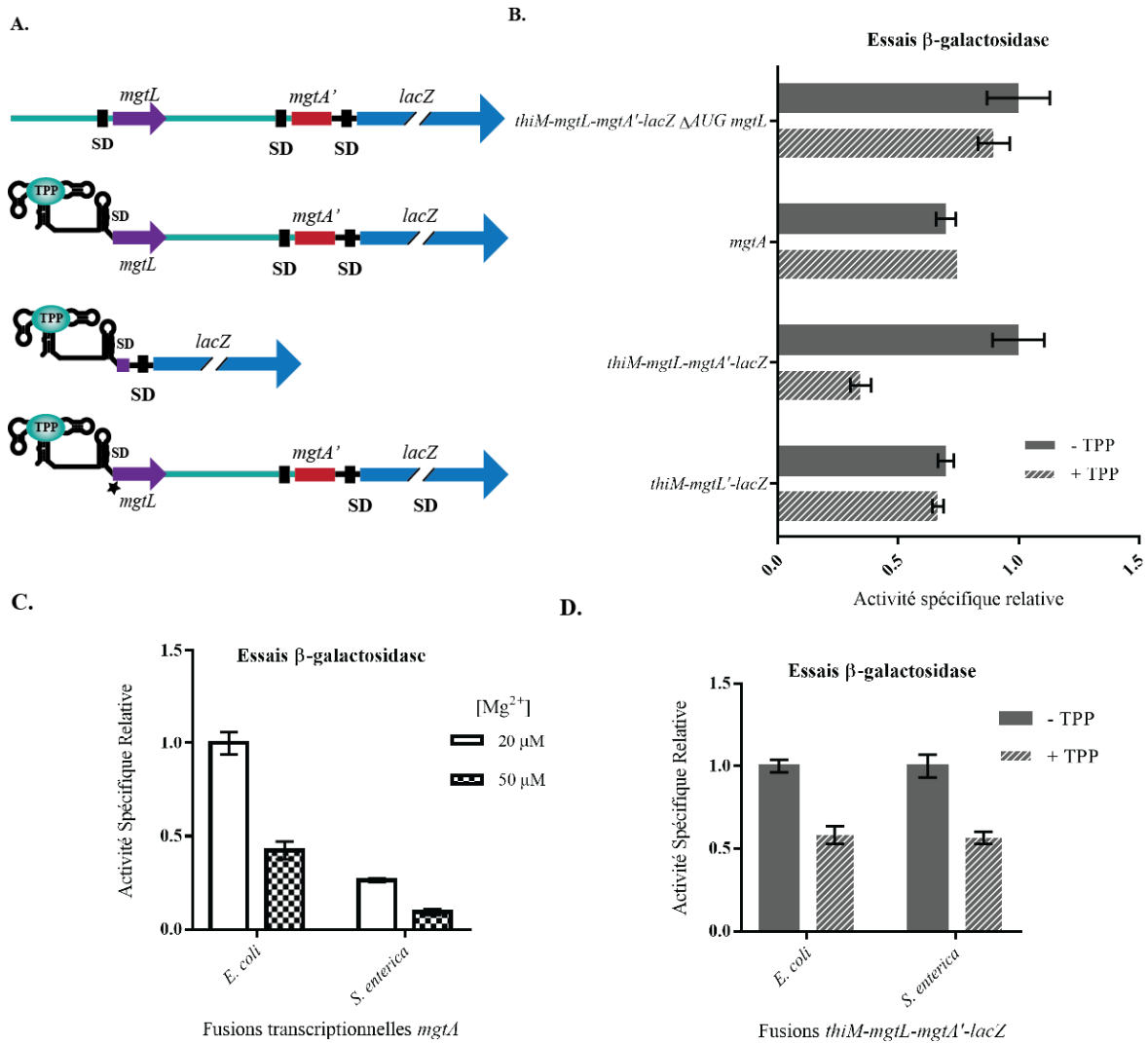


Figure 26. Le contrôle de la traduction de *mgtL* est suffisant pour réguler l'expression de *mgtA*. A) Schémas des fusions utilisées pour les essais β-galactosidase en B. L'étoile indique la mutation de l'AUG de *mgtL* pour un AGG. B) Essais β-galactosidase réalisés sur les fusions transcriptionnelles indiquées en présence ou absence de TPP. C) Essais β-galactosidase réalisés sur des fusions transcriptionnelles *mgtA* de *E. coli* ou *S. enterica* dans *E. coli*. D) Essais β-galactosidase réalisés sur la fusion hybride *thiM-mgtL-mgtA'-lacZ* avec la séquence de *mgtL-mgtA'* de *E. coli* ou *S. enterica* dans *E. coli*. Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

Ces conclusions sont en opposition avec celles provenant des observations faites chez *S. enterica*. Afin de déterminer si les différences observées sont dues à l'organisme étudié, une fusion hybride contenant la séquence de *mgtLA* de *S. enterica* a été construite chez *E. coli*.

Après avoir confirmé que le riborégulateur de *S. enterica* régule l'expression de *mgtA* chez *E. coli* (Figure 26C), les essais β -galactosidase ont été réalisés avec cette fusion et la même répression par le TPP a été observée (Figure 26D). Ces résultats étant en désaccord avec les études précédentes (Gall et al., 2016; Park et al., 2010; Zhao et al., 2011), le même type d'expérience a été reproduit chez *S. enterica*. Pour des raisons techniques, une fusion *gfp* sur plasmide a été construite à la place d'une fusion *lacZ* chromosomique (Figure 27A). Cette fusion a été insérée sur le plasmide pGX10 (Urban et Vogel, 2007) et un immunobuvardage de type *northern* a été réalisé avant et après (5 et 10 min) addition de TPP. La sonde cible spécifiquement la plateforme d'expression de *thiM* d'*E. coli* pour éviter l'hybridation de *thiM* endogène de *S. enterica*. La diminution de l'ARNm observée après addition de TPP avec aussi la séquence d'*E. coli* et de *S. enterica* confirme les résultats obtenus chez *E. coli* (Figure 27B).

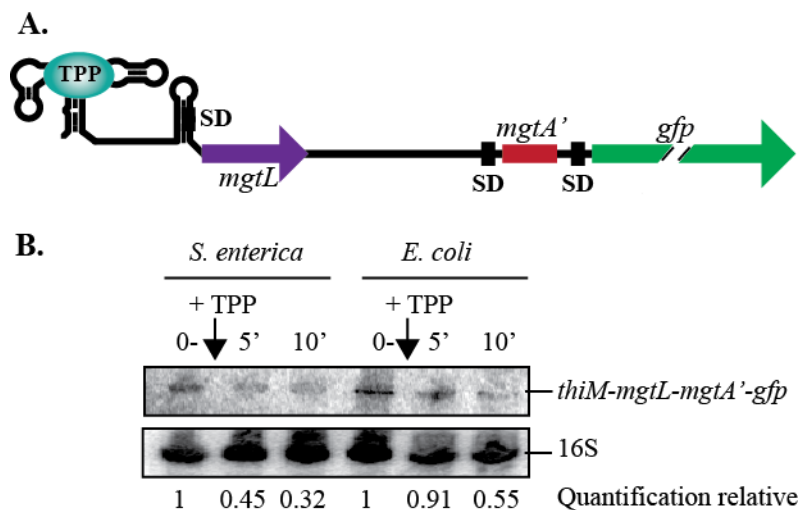


Figure 27. Le peptide leader chez *S. enterica*. A) Schéma de la fusion hybride *thiM-mgtL-mgtA'-gfp*. B) Immunobuvardage de type *northern* avec des ARN extraits de souches de *S. enterica* contenant la construction présenté en A) avec la séquence de *mgtL-mgtA'* de *S. enterica* ou *E. coli*. Les sondes utilisées ciblent la plateforme d'expression du riborégulateur *thiM* de *E. coli* et les ARN 16S. La quantification est relativisée aux 16S.

4.5. LA SÉQUENCE DU PEPTIDE EST IMPORTANTE POUR LE CONTRÔLE DE L'EXPRESSION DE *MGT*A

La différence d'interprétation entre les études précédentes et ces nouvelles données sur la traduction de *mgtL* a été investiguée. La majorité des expériences réalisées dans les études précédentes chez *S. enterica* observait l'effet de l'inhibition de l'élongation de la traduction plutôt que l'initiation comme il a été fait ici. Il est possible qu'en bloquant l'élongation de la traduction, les ribosomes pourraient ne pas être relâchés immédiatement, imitant ainsi un site de pause du ribosome ou *calage*. Plusieurs peptide leader utilisent le *calage* du ribosome pour réguler l'expression du gène en aval, comme pour l'opéron *tna* (Konan et Yanofsky, 1997). Dans le cas de *tnaC*, le *calage* du ribosome au codon de terminaison bloque la liaison de Rho, augmentant ainsi l'expression de *tnaA*. La plupart des travaux sur *mgtL* de *S. enterica* s'appuient sur l'introduction de codons de terminaison prématurés dans le cadre de lecture de *mgtL* (Park et al., 2010; Zhao et al., 2011), ce qui pourrait imiter du *calage*. De plus, l'utilisation d'antibiotiques comme la tétracycline, qui bloque l'attachement des ARNt aminoacylés chargés au site A du ribosome (Chopra et Roberts, 2001; Park et al., 2010), peut également induire le *calage*, autant que des mutants des protéines ribosomiques qui favorisent le *calage* du ribosome sur l'ARNm (Gall et al., 2016). Dans l'ensemble, les résultats obtenus chez *S. enterica* ne sont pas en contradiction avec ceux obtenus ici, mais supportent plutôt un modèle dans lequel le *calage* du ribosome est impliqué. En accord avec ces hypothèses, l'identité de la séquence peptidique est cruciale puisque lorsque celle-ci est remplacée par les 17 premiers codons de *thiM* dans la fusion hybride avec *thiM*, aucune expression n'est observée (Figure 28A et B). Pour investiguer plus loin la possibilité d'un mécanisme de *calage*, la séquence de *mgtL* a été remplacée par une séquence de *calage* connue. Une petite séquence (20 résidus) a été choisie pour imiter au mieux le *calage* potentiel du petit cadre de lecture *mgtL* (17 résidus) (Gollnick et Babitzke, 2002 ; séquence E1). Le niveau d'ARNm obtenu avec cette fusion est ~ 5 fois supérieur qu'avec la séquence de *mgtL*. De plus, la régulation par le TPP peut être observée pour cette séquence, bien que cette régulation soit

plus faible (Figure 28B). Remplacer la séquence de *mgtL* avec une séquence de *calage* restaure avec succès l'expression de *mgtA* perdue avec la séquence de *thiM*, ce qui suggère que *mgtL* pourrait également contenir une séquence de *calage*. Ces résultats expliquent ainsi la différence d'interprétation proposée ici par rapport aux études précédentes et suggèrent un rôle commun de *mgtL* pour les deux organismes.

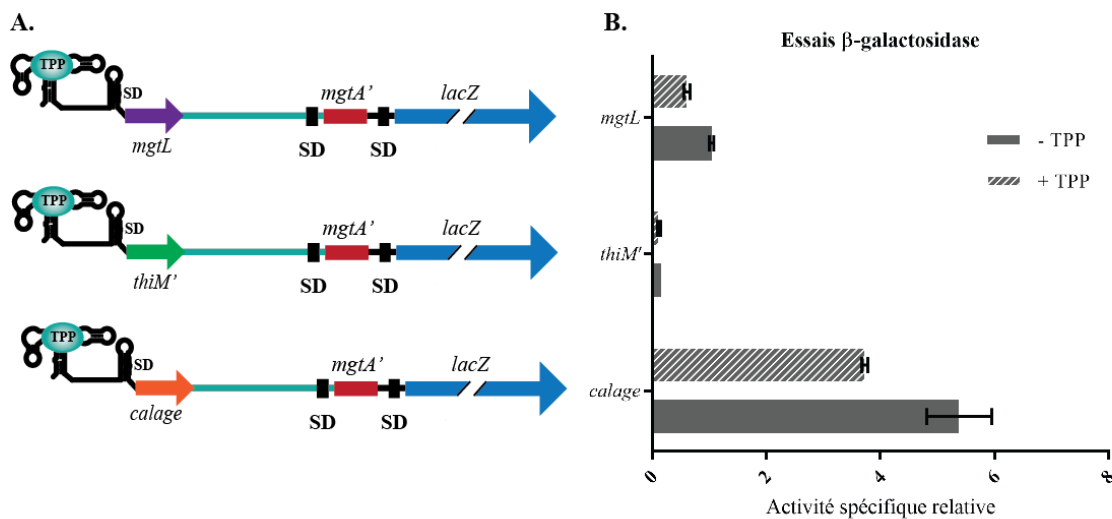


Figure 28. Importance de la séquence du peptide MgtL. A) Schéma des fusions utilisées en B). B) Essais β-galactosidase réalisés avec les fusions indiquées en présence ou absence de TPP. Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

La séquence de MgtL contient quatre prolines et quatre arginines (Figure 29A). Une suite de trois prolines consécutives ou plus est connue pour induire le *calage* du ribosome (Ude et al., 2013). Dans le cas de MgtL, celles-ci ne sont pas consécutives, mais positionnées un résidu sur deux. Une étude récente a étudié l'impact d'une telle disposition de prolines, en prenant l'exemple de MgtL (Chadani et al., 2017). Les auteurs suggèrent que de telles séquences modulent la stabilité des ribosomes, pouvant résulter en leur désassociation de l'ARNm. Bien qu'ils distinguent ce processus du *calage*, leurs résultats confirment que la séquence de MgtL affecte l'élongation de la traduction. Le début de la séquence de MgtL pourrait ainsi induire du *calage* responsable de la régulation de l'expression de *mgtA*. La présence de quatre arginines

plus loin dans la séquence est également intrigante. Plusieurs arginines consécutives ne sont pas connues pour induire du *calage* mais ces codons ne sont pas les plus utilisés chez *E. coli* : leur usage est de 0,35 % pour CGA, 0,54 % pour CGG et 0,20 % pour AGA, contre 2,10 % et 2,21 % pour les plus communs (Spencer et Barral, 2012). L'usage de codons rares ainsi que de codons qui ne possèdent pas d'ARNt complémentaire peut être associé à une diminution locale de la vitesse d'élongation. Ainsi, les codons 3, 7, 9 et 10 n'ont pas d'ARNt complémentaire et les codons prolines ne sont pas les plus abondants. La séquence dans son ensemble pourrait donc ralentir l'élongation de la traduction. Identifier les codons qui pourraient influencer le *calage* semble donc compliqué, d'autant plus que le cadre de lecture se situe dans une structure d'ARN elle-même importante pour la régulation. L'utilisation de la fusion hybride avec *thiM* pourrait donc être nécessaire à l'étude de l'impact des codons. De plus, il est probable que plus d'un codon soit impliqué et que plusieurs mutations soient nécessaires. Cependant, plusieurs éléments pourraient permettre d'identifier le site de *calage* :

- La présence d'une séquence GGAGA, similaire à un SD, au niveau des codons arginines 11 et 12 pourrait maintenir le ribosome à cette position, cet effet pouvant être favorisé par les trois codons rares arginines consécutifs ;
- Un mutant contenant un codon de terminaison prématuré au codon 13 (STOP, 106A) augmente le niveau d'ARNm de *mgtA*, mais conserve une partie de la régulation d'après les essais β -galactosidase (Figure 29B). Le ribosome étant connu pour *caler* à certains codons de terminaison (Konan et Yanofsky, 1997; Wolin et Walter, 1988), l'introduction d'un tel codon à une position proche du site de *calage* pourrait expliquer l'augmentation d'activité observée ;
- De nombreux indices de l'importance de chaîne peptidique naissante sur l'élongation de la traduction indiquent que l'identité de la séquence déjà traduite est importante pour le *calage* (Lu et al., 2007a; Wilson et al., 2016; Woolhead et al., 2006). En particulier, les résidus positivement chargés tels que les arginines et lysines peuvent induire le *calage* (Lu et Deutsch, 2008). De plus, des peptides

antimicrobiens riches en proline peuvent interagir dans le tunnel de sortie du ribosome et inhiber ainsi l'élongation de la traduction (Graf et al., 2017). MgtL contenant de nombreuses prolines, arginines et également une lysine, la combinaison de ces résidus pourrait favoriser le *calage* du ribosome (Figure 29C).

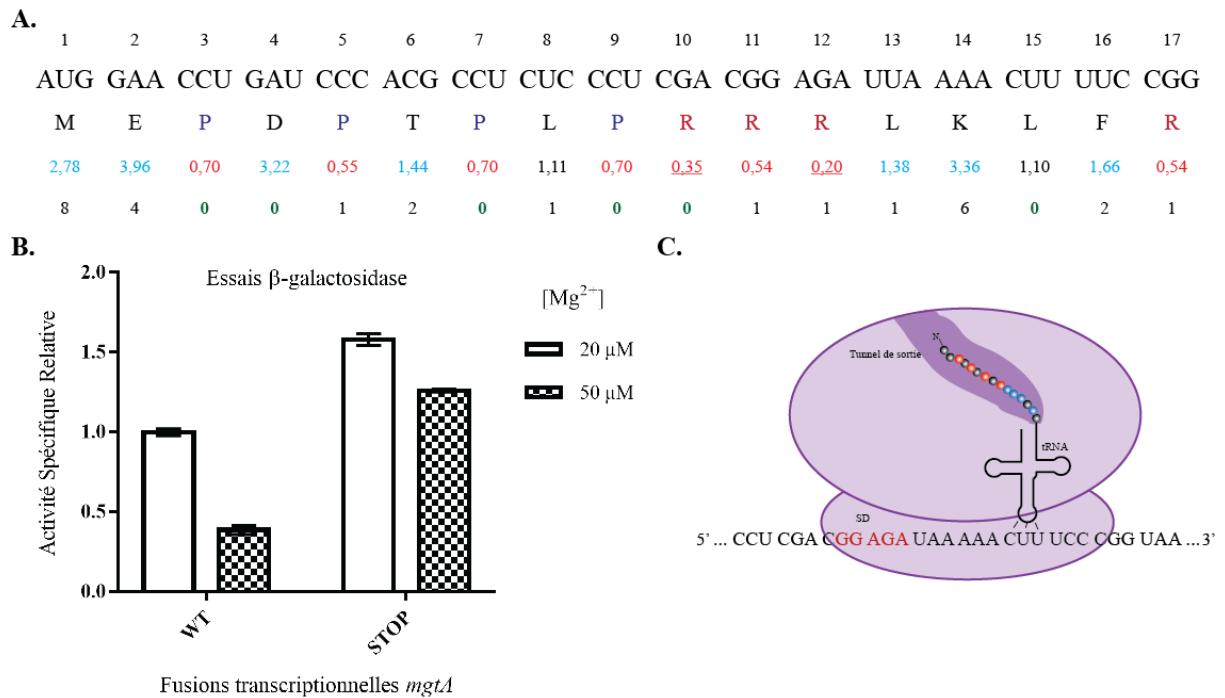


Figure 29. Éléments potentiellement responsables du *calage*. A) Représentation des codons de *mgtL* numérotés avec l'acide aminé correspondant, l'usage et le nombre d'ARNt complémentaires disponibles. Les résidus proline et arginine sont indiqués en bleu et rouge foncé respectivement. Les fréquences d'usage élevées sont indiquées en bleu clair et les fréquences faibles en rouge. Les codons rares sont sous lignés. L'absence d'ARNt complémentaire est indiquée en vert foncé. B) Essais β -galactosidase réalisés avec des fusions transcriptionnelles *mgtA* sauvage (WT) ou mutante (STOP) en faible et fort Mg²⁺. STOP : U106A. C) Modèle proposé du *calage*. La séquence SD est représentée en rouge. Les résidus traduits sont représentés par des billes, rouges pour les prolines, bleues pour les résidus chargés positivement.

4.6. MÉCANISME DE RÉGULATION DU RIBORÉGULATEUR ET DU PEPTIDE LEADER

Les résultats obtenus dans ce chapitre montrent l'importance de la traduction du cadre de lecture MgtL dans la régulation par le Mg^{2+} et donc pour le riborégulateur *mgtA*. L'implication du peptide leader dans la régulation permet d'expliquer la complexité du mécanisme constatée au chapitre 3. La prise en compte de la traduction du riborégulateur est donc nécessaire à l'étude du riborégulateur.

L'importance de la présence des ribosomes peut s'expliquer par la présence d'un site de *calage* du ribosome. Afin de décortiquer le mécanisme, il a été entrepris d'étudier le riborégulateur dans un contexte de traduction *in vitro*. Le kit de traduction PURExpress a été choisi pour sa composition bien définie et parce qu'il contient également le nécessaire pour transcrire, permettant ainsi le couplage transcription-traduction. La polymérase d'ARN de *E. coli* est ajoutée au kit qui contient la polymérase T7 qui ne convient pas à l'étude de la régulation. La concentration de Mg^{2+} étant d'environ 10 mM dans le kit, seule la condition fort Mg^{2+} peut être observée. Néanmoins afin de diminuer artificiellement la concentration de Mg^{2+} , de l'EDTA peut être ajouté pour chélater les ions divalents dont le Mg^{2+} . L'ajout d'EDTA a pour effet une augmentation de l'expression de *mgtA* spécifique puisque le contrôle *lacZ* ne varie pas (Figure 30A). Ces résultats suggèrent que le riborégulateur est fonctionnel dans le kit de traduction *in vitro* et que l'EDTA diminue efficacement la concentration en Mg^{2+} sans affecter la traduction en générale. L'EDTA pouvant toutefois affecter plus que la concentration en Mg^{2+} , la fusion *thiM-mgtL-mgtA-lacZ* a également été étudiée. Le TPP n'a aucun effet sur la traduction de *lacZ* et réprime efficacement la traduction de *thiM in vitro*, mais pas celle de la fusion hybride (Figure 30B). Ce résultat inattendu, cette construction étant capable de réguler *in vivo*, met en évidence une différence majeure entre la traduction *in vitro* et la traduction *in vivo*. Puisque l'EDTA active l'expression de *mgtA in vitro* alors que le TPP ne la réprime pas, il est possible que le ribosome ne puisse pas *caler* dans le kit, optimisé pour produire rapidement et efficacement des protéines. L'ajout d'EDTA et la diminution du Mg^{2+}

pourraient eux favoriser le *calage* autrement moins efficace. Une autre hypothèse est que l'effet observé avec l'EDTA soit purement traductionnel et qu'il n'y ait pas d'effet sur le niveau d'ARNm dû à l'absence d'un facteur transcriptionnel ou une RNase par exemple.

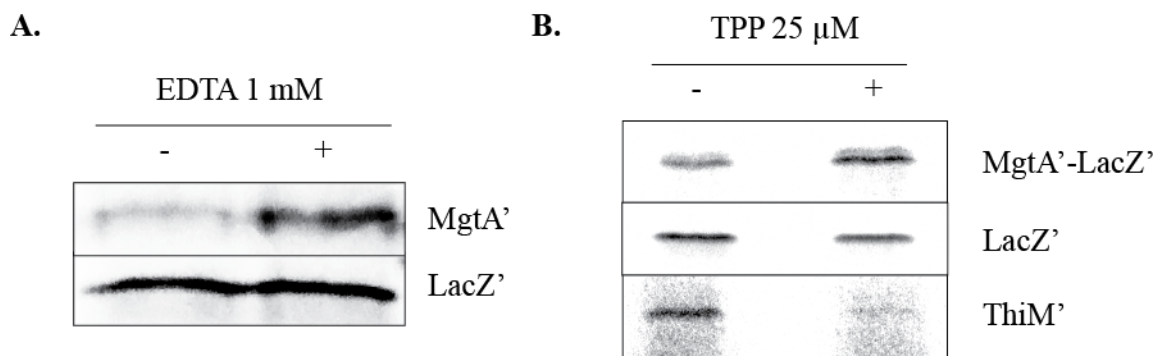


Figure 30. Traduction *in vitro*. A) Transcription-traduction avec le kit PURExpress à partir de matrices d'ADN contenant *PlacUV5* et soit la séquence de *mgtA* à partir du +1 et contenant 260 codons de *mgtA*, soit la séquence 5' du contrôle fourni par le kit et la séquence de *lacZ*. La polymérase de *E. coli* a été ajoutée. Ainsi que de l'EDTA tel qu'indiqué. Les protéines produites à 37°C sont marquées radioactivement grâce à l'ajout de ³⁵S-méthionine. B) Transcription-traduction des protéines indiquées en présence ou absence de TPP tel que mentionné. La protéine de fusion MgtA'-LacZ a été obtenue à partir d'une matrice d'ADN contenant la fusion hybride *thiM-mgtL-mgtA'-lacZ*. Dans cette fusion seule la traduction de MgtA'-LacZ est visualisée sur le gel. La protéine tronquée LacZ est obtenue comme en A. La protéine ThiM est obtenue à partir d'une matrice contenant le riborégulateur et le cadre de lecture thiM.

La possibilité qu'un facteur non identifié soit impliqué a été envisagée étant donné que les diverses expériences conduites *in vitro* ne sont pas concluantes (chapitres 3 et 4). La nécessité de la présence de ribosomes a mené à rechercher un facteur faisant le lien entre la transcription et la traduction. Une étude récente a révélé la structure de l'expressome contenant la polymérase d'ARN et le ribosome, permettant d'expliquer le couplage de la transcription et la traduction et comment un facteur protéique associé au ribosome peut influencer la transcription (Kohler et al., 2017). Dans l'expressome, la protéine ribosomique S10, aussi appelée NusE, se trouve à l'interface 30S et polymérase d'ARN. De plus NusE est capable de lier le facteur de transcription NusG afin d'empêcher la protéine Rho de terminer la transcription des transcrits en cours de traduction (Burmam et al., 2010). Bien que Rho ne

semble pas impliqué dans la régulation, il est possible que ces facteurs jouent un rôle dans le mécanisme de régulation de *mgtA*, expliquant l'importance de la traduction de la région régulatrice. NusE et NusA sont essentiels à la survie de *E. coli* (Sen et al., 2008), rendant impossible l'utilisation de mutants pour étudier leur impact dans la régulation. Cependant le facteur NusB semble important pour le chargement de NusE et un mutant viable de ce dernier est disponible. Malheureusement ce mutant ne se multiplie pas suffisamment avec les conditions utilisées dans cette étude pour obtenir des résultats concluants. Un autre facteur semblable à NusG, RfaH, a également été montré pour lier à la fois la polymérase d'ARN et NusE (Knauer et al., 2012). Un mutant étant disponible, l'impact de RfaH a été évalué par essais β -galactosidase avec la fusion transcriptionnelle *mgtA*. La mutation n'a qu'un effet limité sur l'expression ou la régulation de *mgtA* (Figure 31). Ce résultat n'est pas surprenant puisque l'action de RfaH nécessite la présence d'une séquence *ops* (*operon polarity suppressor*) pour agir sur la transcription dans les cas observés jusqu'à présent, séquence absente de la région 5' de *mgtA*.

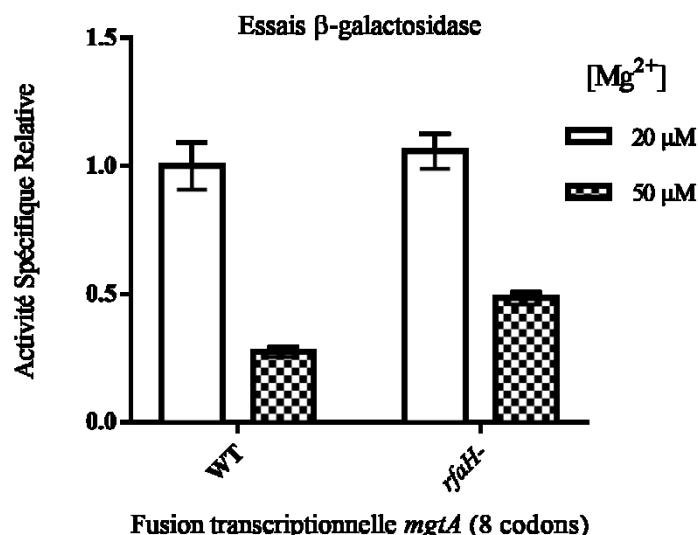


Figure 31. RfaH n'est pas impliqué dans le mécanisme de régulation. Essais β-galactosidase réalisés sur les souches sauvage (WT) ou mutante (*rfaH*) contenant la fusion transcriptionnelle *mgtA*.

Une des raisons avancées pour la nécessité du couplage entre la transcription et la traduction est la prévention des *R-loops* (Gowrishankar et Harinarayanan, 2004). Chez *E. coli*, les *R-loops* sont défaites par la RNase H qui dégrade les ARN impliqués dans des hybrides ADN:ARN (Aguilera et García-Muse, 2012). La formation de *R-loops* a également été associée avec la régulation de l'expression génique pour avoir un rôle dans la cellule (Wahba et Koshland, 2013) et pourrait grâce à la RNase H constituer un mécanisme alternatif pour la dégradation des ARNm chez *E. coli* (Anupama et al., 2011). Ainsi la traduction de *mgtL* pourrait empêcher la formation d'une *R-loop* normalement défaite par la RNase H. Dans le but de tester cette hypothèse, un mutant de délétion *rnhA*⁻ et un plasmide de surexpression ont été utilisés pour évaluer l'effet de l'absence de RNase H et celle de sa surexpression. Les essais β-galactosidase montrent que ni la délétion ni la surexpression n'ont d'effet sur l'expression ou la régulation de *mgtA*, suggérant que la RNase H ne joue pas un rôle dans le mécanisme de régulation (Figure 32A et B).

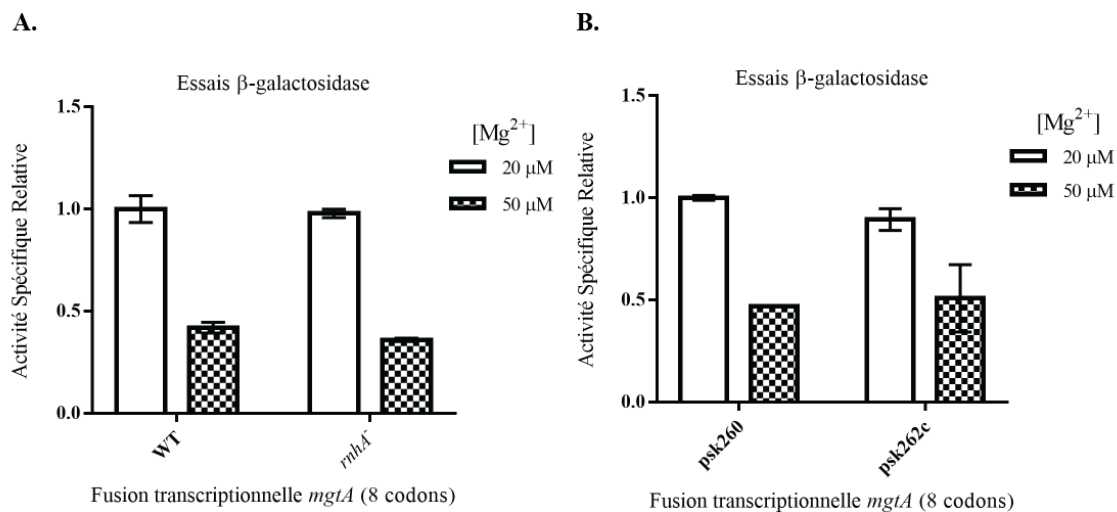


Figure 32. La RNase H n'est pas impliquée dans le mécanisme de régulation. A) Essais β -galactosidase réalisés avec les souches sauvage (WT) et mutante (*rnhA*) contenant la fusion transcriptionnelle. B) Essais β -galactosidase réalisés avec la souche contenant la fusion transcriptionnelle *mgtA* et le plasmide indiqué.

Récemment une étude a montré que la petite protéine MgtS (précédemment YneM) régule l'activité de la protéine MgtA afin d'augmenter la concentration intracellulaire en Mg^{2+} (Wang et al., 2017). Pour évaluer les niveaux de Mg^{2+} intracellulaire, les auteurs ont utilisé une fusion de la région 5' de *mgtA* à *lacZ*. Ils ont alors montré que la délétion de MgtS activait l'activité de cette fusion, concluant que cette activation était due à la diminution de la concentration en Mg^{2+} . Comme aucun contrôle n'a montré que MgtS ne pouvait affecter le niveau d'ARNm de *mgtA*, cette hypothèse a été testée ici. Pour cela, une délétion de *mgtS* a été faite dans la souche contenant la fusion hybride *thiM-mgtL-mgtA*. Les essais β -galactosidase montrent que la délétion de *mgtS* n'affecte pas la régulation par le TPP, suggérant que MgtS affecte bien la concentration intracellulaire en Mg^{2+} et non pas le niveau d'ARNm (Figure 33).

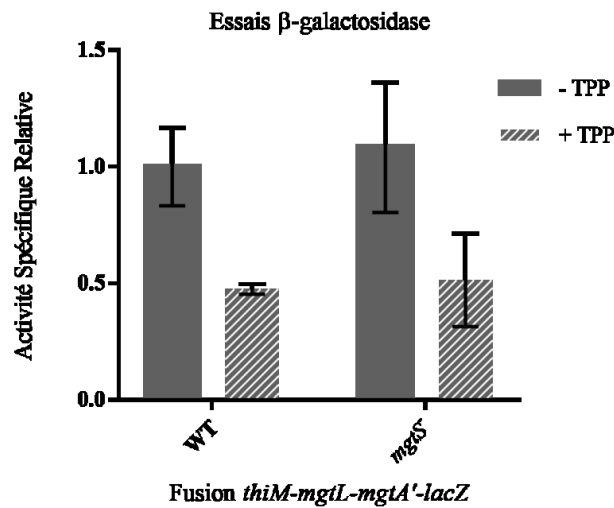


Figure 33. MgtS n'est pas impliqué dans le mécanisme de régulation. Essais β -galactosidase réalisés avec les souches sauvage (WT) ou mutante (*mgtS*) contenant la fusion transcriptionnelle .

La nécessité de la traduction de la première partie du riborégulateur et l'absence d'information sur le rôle exact de la deuxième structure ont amené à investiguer quelles parties de la région 5' étaient nécessaires à la régulation. Des fusions courtes, contenant soit la région 5', soit la région 3' du riborégulateur ont été construites (Figure 34A et C). Les essais β -galactosidase montrent que la région contenant le peptide leader n'est pas suffisante pour la régulation et que la deuxième structure est probablement requise au complet (Figure 34B). La deuxième structure n'est pas non plus suffisante pour la régulation (Figure 34D). Ces résultats suggèrent que la présence de ribosomes sur *mgtL* influence l'élongation de la transcription au niveau de la deuxième structure.

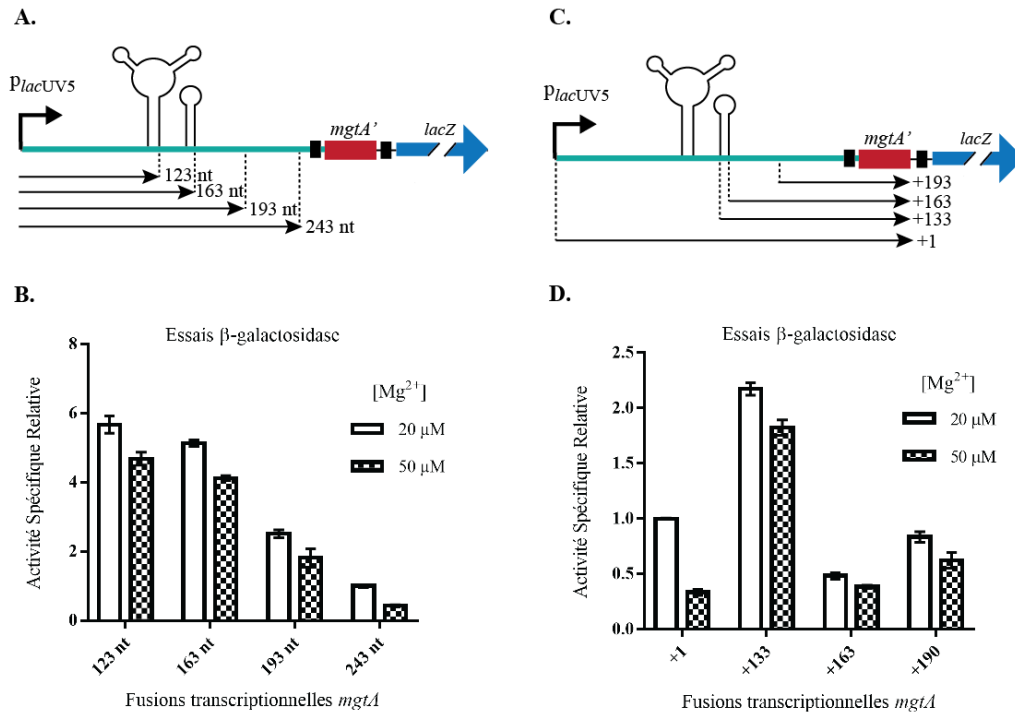


Figure 34. Séquences requises pour la régulation. A) Schémas des fusions transcriptionnelles raccourcies en 3'. B) Essais β-galactosidase réalisées avec les fusions présentées en A). C) Schémas des fusions transcriptionnelles raccourcies en 5' et qui contiennent les 8 premiers codons de *mgtA*. D) Essais β-galactosidase réalisées avec les fusions présentées en C). Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

L'ensemble des résultats obtenus aux chapitres 3 et 4 permettent d'établir un modèle du mécanisme de régulation employé par le riborégulateur *mgtA* chez *E. coli*. La concentration intracellulaire en Mg^{2+} libre détermine la conformation adoptée par la partie 5' du riborégulateur (Figure 35, 1). La structure adoptée détermine alors l'accessibilité de la séquence SD du peptide leader *mgtL*, et donc l'initiation de sa traduction, plus forte en faible Mg^{2+} (Figure 35, 2). La présence de ribosomes, renforcée par le *calage* en faible Mg^{2+} , détermine alors le sort de la polymérase d'ARN pausé au C214. En absence de ribosomes, en fort Mg^{2+} , la polymérase décroche de l'ARNm, terminant ainsi la transcription. À l'inverse, en présence de ribosomes en faible Mg^{2+} , la polymérase continue l'élongation de la transcription, permettant l'expression de *mgtA* (Figure 35, 3 et 4).

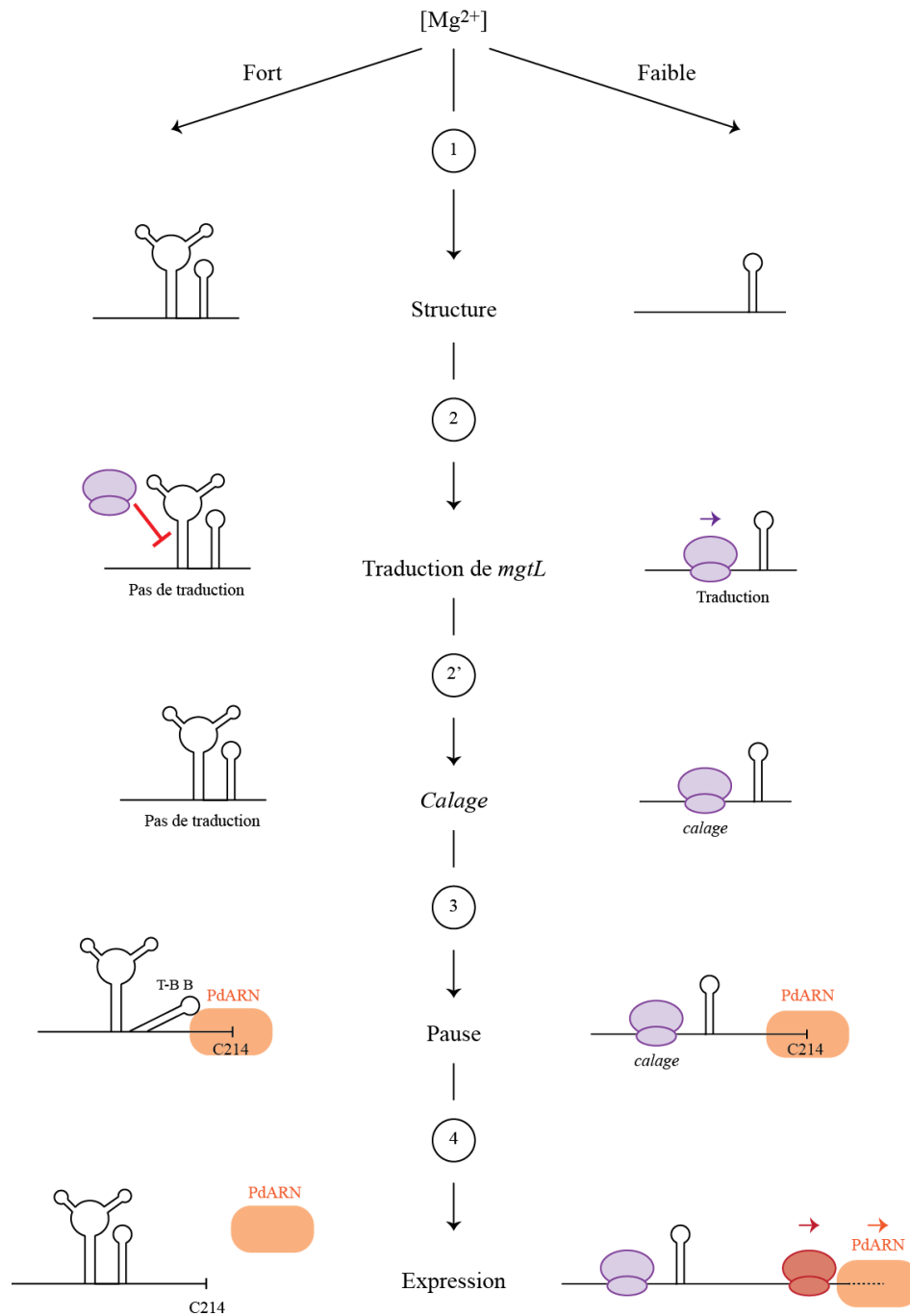


Figure 35. Modèle de la régulation par le riborégulateur et le peptide leader. La concentration en Mg^{2+} détermine la structure adoptée par la région 3' du riborégulateur (1). La structure contrôle l'initiation de la traduction de *mgtL* (2). Les ribosomes traduisant *mgtL* calent sur l'ARNm (2'). La présence de ribosomes sur l'ARNm permet d'empêcher une structure de renforcer la pause au C214 (3). La polymérase d'ARN n'est capable de reprendre l'élongation de la transcription qu'en faible Mg^{2+} et décroche suite à la pause en fort Mg^{2+} . La transcription du transcrit pleine-longueur permet aux ribosomes de traduire *mgtA*.

Le mécanisme exact par lequel la présence de ribosome influence la polymérase à la pause reste à déterminer. Néanmoins, comme la traduction d'autres peptides leaders a été montrée pour stabiliser des structures requises pour la terminaison ou l'antiterminaison (Gollnick et Babitzke, 2002) et comme une étude chez *S. enterica* suggère que la tige-boucle B interagit avec la polymérase pour renforcer la pause (Hollands et al., 2014), il est possible que la présence de ribosomes sur *mgtL* empêche la formation de la tige-boucle B pour permettre à la polymérase de repartir et poursuivre la transcription. La structure active du riborégulateur pourrait être trop instable en comparaison de la structure inactive, expliquant à la fois la nécessité de la présence de ribosome ainsi que la difficulté de déterminer la structure active *in vitro*.

CHAPITRE 5

ANALYSE DES DONNÉES DE PROFILAGE RIBOSOMIQUE

5.1. LE PROFILAGE RIBOSOMIQUE POUR ÉVALUER L'EXPRESSION GÉNIQUE.

La majorité des études visant à évaluer l'expression génique, évalue en fait la quantité d'ARNm par séquençage ARN haut débit (ou puces à ADN), et ne tient donc seulement compte que de l'abondance de transcrits. Elles ne prennent ainsi pas en compte la possibilité que l'ARNm ne soit pas traduit ou de la régulation traductionnelle dans le cas d'une étude d'expression différentielle. Il est ainsi impossible par RNA-seq d'observer la régulation traductionnelle effectuée par un petit ARN tel que celle de RyhB sur *yegD* (Wang et al., 2015).

Une technique relativement récente a été mise au point afin de pallier cette faiblesse : le profilage ribosomique ou *ribosome profiling* (Ingolia et al., 2009). Cette technique consiste à extraire les ARNm totaux, à digérer par la suite les fragments d'ARN non protégés par les ribosomes, à purifier les monosomes obtenus puis à séquencer ces 'empreintes' ou *footprints* de ribosomes (

Figure 36). La synthèse protéique est ainsi prise en compte en plus du niveau de transcription.

Généralement un séquençage ARN classique est réalisé en parallèle afin d'étudier le taux de traduction, ce qui permet de différencier si un changement d'expression provient d'une variation de la quantité d'ARNm ou de l'efficacité de traduction. Depuis la première étude par Ingolia *et al.* (2009), cette approche a permis :

- De mettre en évidence de nouveaux cadres de lectures (Mumtaz et Couso, 2015; Oh et al., 2011).
- D'identifier les sites de pause du ribosome (Li et al., 2012; Woolstenhulme et al., 2015).
- D'identifier de nouvelles cibles de sRNA (Wang et al., 2015).
- D'étudier l'impact de protéines naissantes sur la traduction (Oh et al., 2011).

Dans le chapitre 4, la technique de *ribosome profiling* a été utilisée pour observer directement l'expression d'un peptide leader et du gène en aval, sans utiliser de fusion ou d'étiquette (ou tag) qui ont pour effet de modifier la structure de l'ARNm dans cette région. Elle a ainsi permis de mettre en évidence la régulation de l'expression de *mgtL* par la concentration en Mg^{2+} . Cependant, de nombreuses autres données ont été obtenues concernant le reste du génome. Le but de ce chapitre est d'extraire et d'analyser un maximum d'informations de ces données.

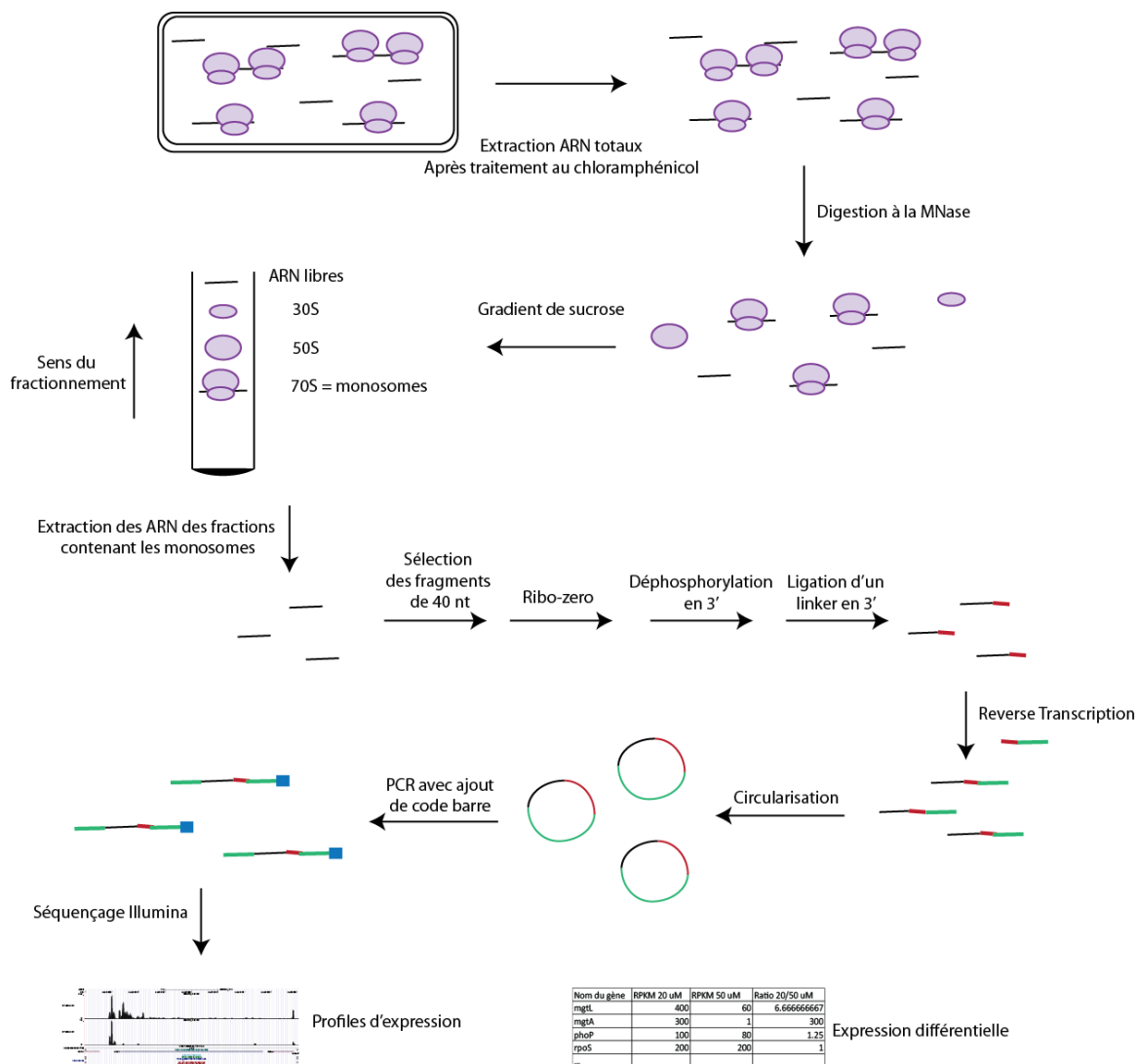


Figure 36. Principe de la technique de profilage ribosomique. Le chloramphénicol est utilisé pour immobiliser les ribosomes sur l'ARN. Suite à l'extraction, les polysomes sont digérés à la nucléase microccale (MNase) qui ne digère pas les fragments d'ARN liés aux ribosomes. Les monosomes obtenus sont purifiés sur un gradient de sucrose. Les petits fragments d'ARN sont ensuite extraits et les fragments entre 20 et 40 nucléotides sont sélectionnés sur gel d'acrylamide. Une banque d'ADNc est réalisée par la suite avec ajout de code barre pour différencier les différents échantillons. Suite au séquençage Illumina, les lectures ou reads sont alignés contre le génome et le niveau relatif d'expression est calculé.

5.2. QUALITÉ DES DONNÉES

Pour les besoins de l'étude du riborégulateur *mgtA*, le *ribosome profiling* a été réalisé lors de la phase de transition vers la phase stationnaire. Pendant cette phase de croissance très courte, beaucoup de changements physiologiques s'opèrent et se traduisent par des variations de l'expression génique (Chang et al., 2002). Étant donné le faible nombre de réplicats (duplicata pour chaque condition), la qualité des données a été étudiée de manière approfondie.

Après alignements des *reads* contre le génome d'*E. coli*, le nombre de lectures de séquençage alignés pour chaque gène et chaque région intergénique pour chaque échantillon a été calculé.

Afin de vérifier que la distribution des *reads* est équivalente dans les différents échantillons, un histogramme a été tracé. Chaque barre correspond au nombre de *reads* (comptes ou *counts*) (en ordonnée) pour un nombre de gènes (en abscisse). Un histogramme a été tracé pour chaque échantillon (Figure 37A). Dans le but de mieux visualiser les données, il est nécessaire de transformer les valeurs brutes. Le log en base 2 est généralement utilisé puisqu'il facilite la reconversion dans l'échelle originale : une différence de 1 en log base 2 correspond à un ratio de 2 sur l'échelle originale. Étant donné que la valeur de compte peut être de 0 dans certaines conditions, il est plus judicieux d'utiliser des « pseudo-comptes », en transformant les données brutes sous la forme : $y = \log_2(\text{comptes} + 1)$. Les histogrammes sont alors plus faciles à interpréter (Figure 37B). La distribution de chacun des échantillons est semblable, à l'exception des gènes les plus exprimés qui correspondent probablement à de la contamination d'ARNr ou ARNt (l'analyse des données aberrantes sera discutée plus tard).

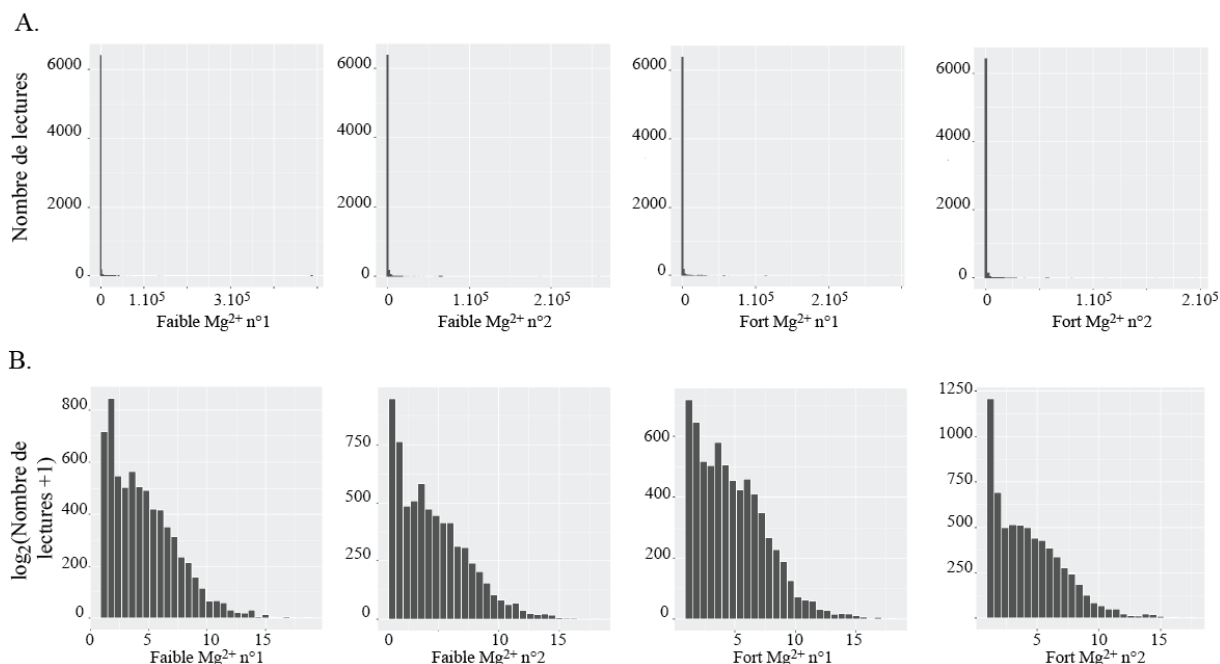


Figure 37. Distribution du nombre de lectures pour les échantillons de profilage ribosomique. Répartition du nombre de lecture par nombre de gènes pour les données brutes A) ou transformées en \log_2 B).

Une autre méthode de comparaison des échantillons est le tracé de boîtes à moustaches ou boîte de Tukey qui permet de résumer les principales caractéristiques statistiques d'une population (médiane (rétrécissement), quartiles (hauteur de la boîte), 1,5 IQR (les « moustaches » de longueur 1,5 fois l'écart interquartile) et taille de l'échantillon (largeur de la boîte)). L'utilisation de boîtes « entaillées » ou *notched* permet de comparer des populations : cette représentation permet de visualiser un intervalle de confiance à 95% autour de la médiane. Les points où la boîte se resserre représentent les bornes de cet intervalle. Deux populations pour lesquelles le resserrement ne se recoupe pas sont généralement considérées comme différentes. Les résultats obtenus (Figure 38) semblent indiquer de légères différences bien que les boîtes montrent une tendance commune. On observe une petite différence entre l'échantillon faible Mg^{2+} n° 1 et le n°2 et entre le faible Mg^{2+} n°1 et le fort Mg^{2+} n°1 mais pas entre le faible n°2 et le fort n°2. Les résultats suggèrent qu'il y a plus un effet entre les 2 réplicats plutôt qu'un effet Mg^{2+} : les boîtes n°2 pour les deux conditions sont légèrement plus

basses que celles du réplicat n°1. Cette différence peut s'expliquer par le fait que les deux réplicats ont été réalisés deux jours différents et que la phase de transition est courte et très variable.

Les points observés au-dessus des « moustaches » représentent des valeurs extrêmes ou atypiques. Les régions les plus abondantes (67%, Figure 39) correspondent à des ARNt qui sont co-purifiés avec le ribosome puisqu'ils sont présents lors de la traduction. Des ARNr sont également présents malgré l'étape de *Ribo-Zero* effectuée pour les éliminer avant le séquençage. Selon Illumina, le *Ribo-Zero* peut enlever 99 % des longs ARNr et 85 % des courts. Les ARNr aux comptes élevés correspondent aux ARN 23S des 7 opérons. Les 23S étant de longs ARN ils auraient dû être retirés plus efficacement. La visualisation de ces ARNr avec *Genome Browser* permet de mettre en évidence une petite région d'environ 30 nucléotides qui est sur-représentée par rapport au reste de l'ARNr. En plus des composants du ribosome, des ARNm sont également retrouvés parmi ces valeurs extrêmes. Une partie de ces ARNm codent pour des protéines ribosomiques. Les ARNm des protéines considérées comme les plus abondantes (*tufA*, *tufB* et *lpp*) se retrouvent également parmi les valeurs extrêmes. Ces dernières ne sont ainsi pas des valeurs aberrantes, mais représentative de l'expression génique.

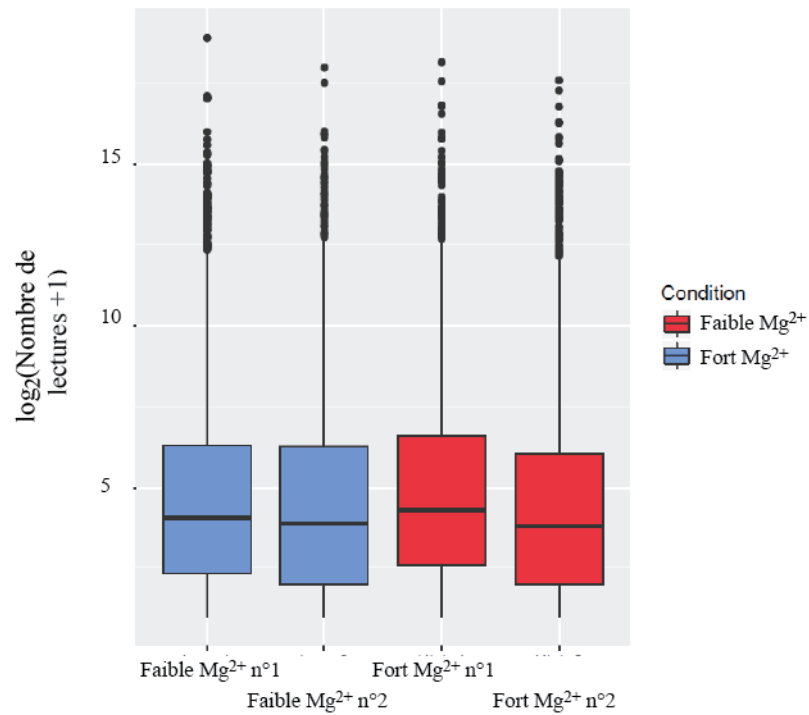


Figure 38. Distribution du nombre de lectures en boîtes à moustaches. La boîte contient la moitié des valeurs, la limite basse représente la valeur du premier quartile Q1, la limite haute celle du troisième Q3. La barre du milieu correspond à la médiane. Les moustaches sont de taille 1,5 fois l'écart interquartile Q3-Q1. Les données à l'extérieur des moustaches sont des données « aberrantes » représentées par des points.

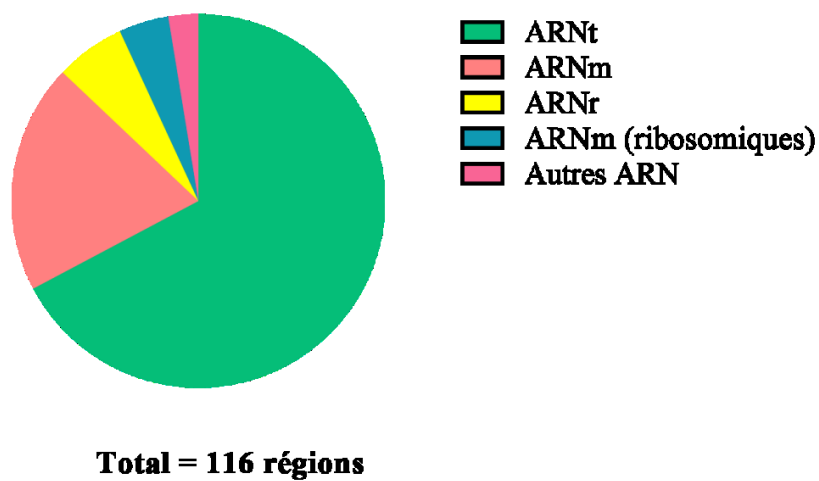


Figure 39. Répartition des ARN selon leur fonction.

Une partie des gènes identifiés ci-dessus ont des valeurs extrêmes que pour un ou deux échantillons. Par exemple, *zinT* est très exprimé dans les échantillons du réplicat n°1 et très peu dans le 2^{ème} (Figure 40). Précédemment nommé *yodA*, *zinT* code pour une protéine supposée être un facteur de stress général capable de lier divers ions métalliques tels que le cadmium, le zinc et le nickel (David et al., 2003). Son expression est activée par une carence en zinc (Graham et al., 2009). Les résultats du *ribosome profiling* indiquent ainsi que la disponibilité en zinc a varié entre les deux réplicats. Il y a probablement eu une contamination en zinc le 2^{ème} jour, inhibant ainsi l'expression de *zinT*, le milieu minimum utilisé ne contenant pas de zinc.

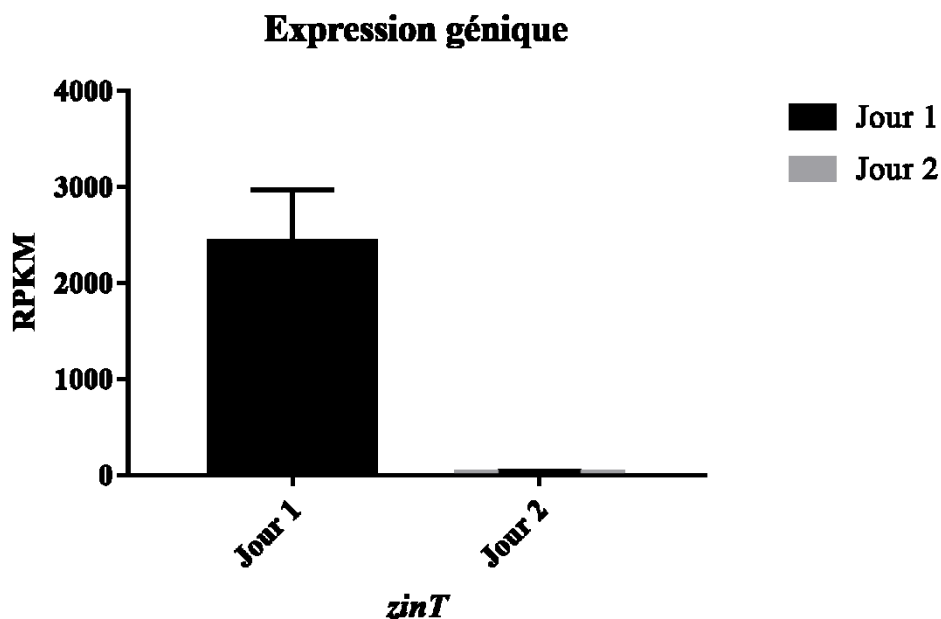


Figure 40. Expression de *zinT*. Résultats du profilage ribosomique pour le gène *zinT*. RPKM : *read per kilobase per million*. Les résultats représentent la moyenne et l'écart-type des deux échantillons du jour 1 et du jour 2.

L'étude de la reproductibilité intra-condition au-delà de la dispersion semble donc nécessaire. Le facteur de corrélation linéaire de Pearson de l'expression des gènes a ainsi été calculé pour

les deux conditions de Mg^{2+} utilisées (Figure 41). Les coefficients étant élevés, 0,957 et 0,962 respectivement, les résultats sont reproductibles dans l'ensemble.

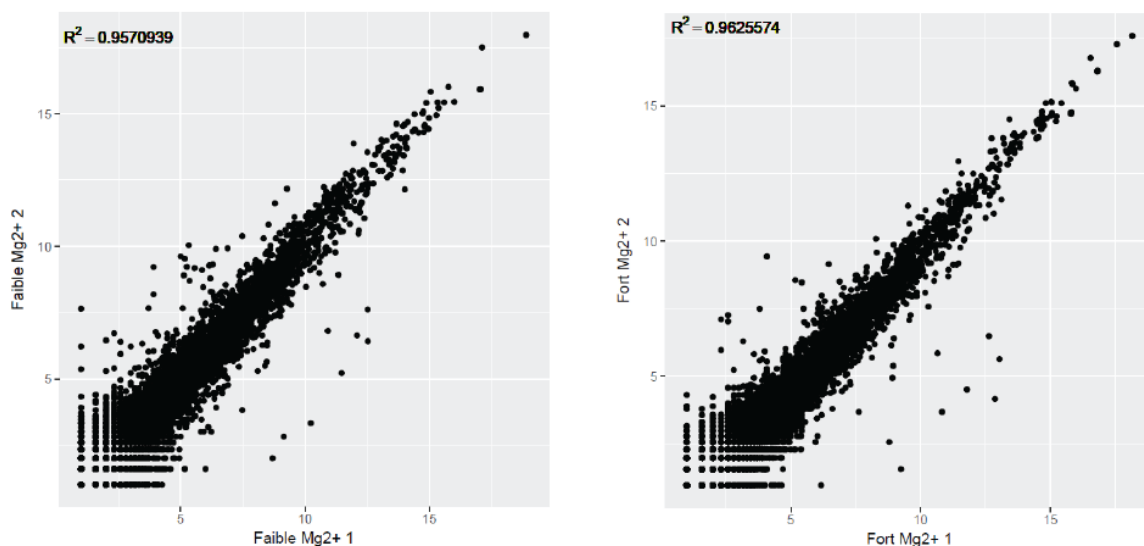


Figure 41. Comparaison de la reproductibilité intra-condition. Chaque point représente un gène avec son nombre de lecture en \log_2 pour chaque jour. Le coefficient de corrélation linéaire de Pearson est indiqué en haut à gauche.

Une autre méthode pour évaluer la reproductibilité est de tracer des MA-plot (log du ratio en fonction du log moyen). Ainsi, dans un MA-plot, les gènes avec un niveau d'expression similaire dans deux échantillons se trouvent autour de la ligne horizontale $y = 0$. Un MA-plot a été tracé pour chacune des conditions (Figure 42 A et B). La ligne rouge représente une régression locale (LOESS) qui permet de mettre en évidence des biais dans le niveau d'expression moyen. Chaque gène est représenté par un cercle semi-transparent pour plus de clarté. Les résultats suggèrent que les données sont moins reproductibles pour la condition fort Mg^{2+} . Le nombre total de lecture relativement différent (4,6 et 3,8 millions de lectures) pouvant être en cause, de nouveaux graphiques ont été tracés avec les données normalisées au nombre total de lectures (Figure 42 C et D). La courbe de régression locale se rapproche de la ligne $y = 0$ pour la condition fort Mg^{2+} , mais ne change pas pour la condition faible Mg^{2+} ce

qui suggère que les divergences observées précédemment sont en partie dues à la taille de l'échantillon et qu'il sera donc important pour la suite de ne pas la négliger.

De manière générale, les résultats semblent reproductibles malgré la phase de croissance étudiée et le faible nombre de réplicats : les données intra-conditions corrélaient bien et les gènes connus pour être extrêmement bien exprimés possèdent un nombre de lectures également élevé. La prochaine étape de validation des données consiste à vérifier que des gènes connus pour être régulés par le Mg^{2+} le sont effectivement dans cette expérience.

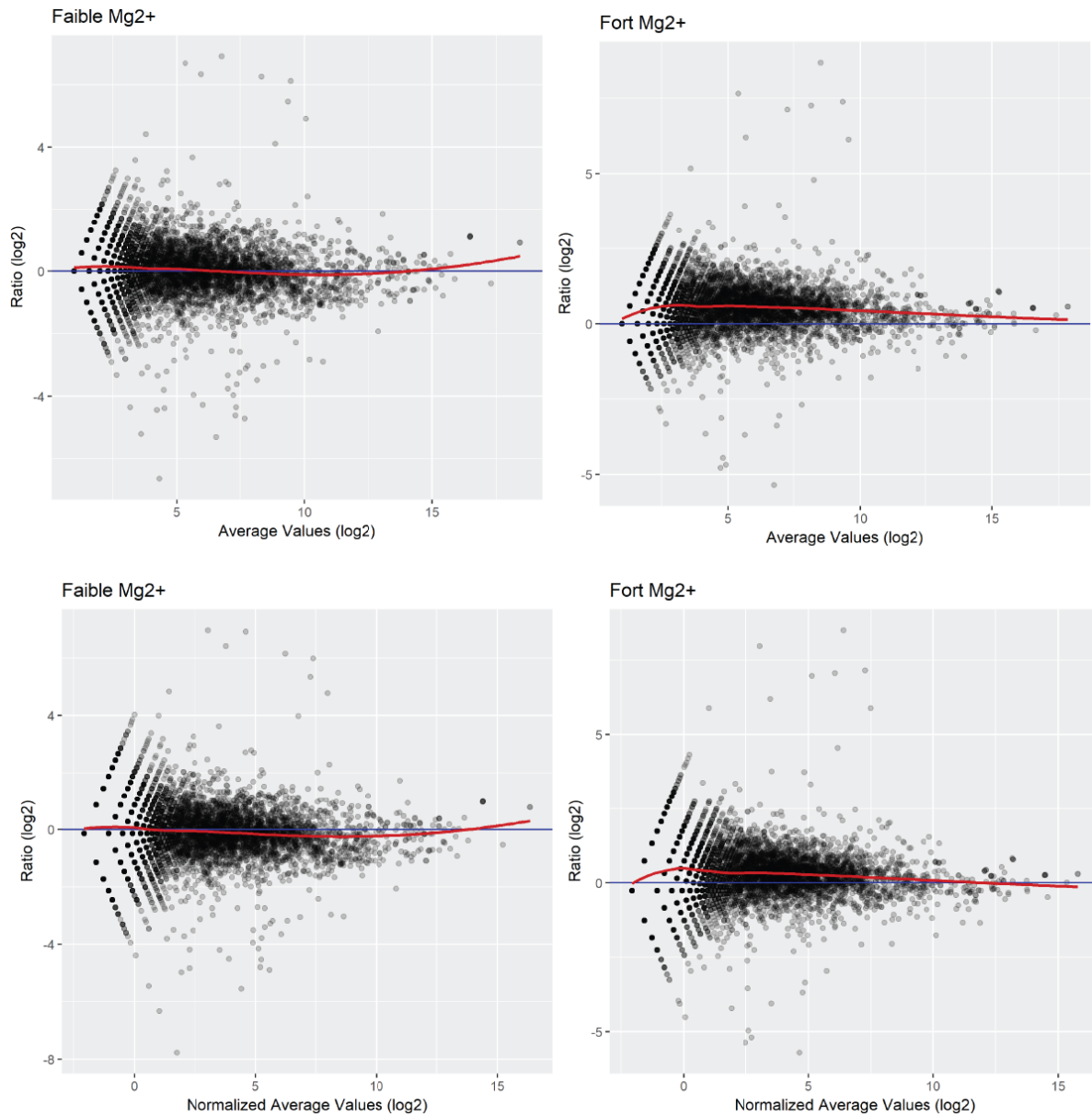


Figure 42. MA-plot intra-condition avant et après normalisation à la taille de l'échantillon. A) MA-plot en faible Mg²⁺ sans normalisation. B) MA-plot en fort Mg²⁺ sans normalisation. C) MA-plot en faible Mg²⁺ avec normalisation. D) MA-plot en fort Mg²⁺ avec normalisation. La courbe en rouge représente une régression locale.

5.3. EXPRESSION DIFFÉRENTIELLE DES GÈNES RÉGULÉS PAR LE Mg^{2+}

Afin d'évaluer l'effet du Mg^{2+} sur l'expression génique, l'expression différentielle des gènes a été étudiée grâce à la suite bio-informatique R DESeq2 (Love et al., 2014). Le principe repose sur le test d'hypothèses :

- H_0 : le gène n'est pas différentiellement exprimé ;
- H_1 : le gène est différentiellement exprimé entre les deux conditions.

L'utilisation d'un test statistique permet de déterminer à quel point l'hypothèse nulle H_0 est possible. Le résultat du test est converti en valeur-p qui est la probabilité de voir un résultat aussi extrême ou plus extrême que la donnée observée, lorsque l'hypothèse nulle est vraie. Ainsi une valeur-p élevée indique qu'il est très probable que H_0 soit vraie. Il y a deux types d'erreurs possibles avec l'interprétation des résultats statistiques : l'erreur de type I (ou faux-positif) est le cas où l'hypothèse nulle est vraie alors que le test statistique suggérerait qu'elle était fausse. Ce type d'erreur est potentiellement dangereux puisqu'elle peut amener à une fausse découverte ; l'erreur de type II (ou faux-négatif) est le cas inverse, moins dangereux que le précédent, mais à éviter pour une détection efficace des gènes différentiellement exprimés. Dans les tests d'hypothèses, la probabilité « acceptable » de commettre une erreur de type I est conventionnellement 0,05. Des problèmes surviennent lorsque plus d'un test d'hypothèse est réalisé puisque chaque test a une probabilité de produire une erreur de type I. Les erreurs s'accumulent ainsi lors de tests multiples (*multiple testing*). Il existe ainsi des méthodes de correction pour contrôler le nombre d'erreurs de type I. Le paquet DESeq2 utilise la correction de Benjamini & Hochberg (Benjamini et Hochberg, 1995) et donne les valeur-p ajustées correspondantes. Pour confirmer la validité des résultats, il a été vérifié que (1) des gènes de maintenance non régulés par le Mg^{2+} l'étaient bien dans cet essai et que (2) les gènes du *stimulon* Mg^{2+} étaient bien régulés. Les résultats sont présentés sur la Figure 43. La

sélection des gènes de maintenance est inspirée de celle des analyses de séquençage ARN et de PCR quantitative, à l'exception des gènes ribosomiques non traduits et des gènes peu exprimés dans les conditions de l'essai. Les gènes *rssA* et *ihfB* sont des gènes de références utilisés pour normaliser des données de qPCR (Zhou et al., 2011) (les autres gènes utilisés sont trop peu exprimés ici). Des gènes impliqués dans plusieurs fonctions de maintenance ont également été choisis : *rpoA*, *rpoS* et *rpoD* sont impliqués dans la transcription, *gyrB* dans la réplacation de l'ADN, *hns* code pour un facteur de régulation transcriptionnelle global et *groL* pour une chaperonne. Ces gènes présentent des ratios (en log, base 2) proches de 0, donc dont l'expression varie peu ou pas avec le Mg^{2+} (Panneau de gauche sur la Figure 43). Il est intéressant de noter que *rpoS* est plus exprimé que *rpoD* (valeurs respectives) : *rpoS* code pour le facteur σ^S associé avec la phase stationnaire et *rpoD* pour le facteur σ^{70} associé avec la phase exponentielle de croissance. Ces résultats confirment que les cultures bactériennes à partir desquelles ont été extraient les *footprints* de ribosome étaient dans le début de la phase stationnaire, puisque *rpoD* est toujours un peu exprimé. Le facteur de régulation PhoP régule environ 5 % des gènes chez *S. enterica* (directement ou indirectement) (Harari et al., 2010). Les gènes contenant des PhoP box (séquence de liaison à PhoP au niveau de la région promotrice) sont ainsi nombreux. Ces gènes ne sont pas tous exprimés et régulés dans cet essai, mais 39 % des gènes trouvés différentiellement exprimés sont connus pour être régulés par PhoP. Le *stimulon* Mg^{2+} tel que décrit en 2003 (Minagawa et al. 2003) comprend 9 gènes possédant au moins une boîte PhoP. Parmi eux, *mgrB*, *rstA*, *slyB*, *phoP* et *borD* (anciennement *vboR*) sont significativement régulés ; *yrbL*, *nagA* et *hemL* sont régulés, mais avec une valeur-p ajustée > 0.1 (Figure 43, panneau du centre). Le dernier gène est *mgtA* dont l'expression n'est pas régulée par PhoP dans la souche utilisée. Les autres gènes PhoP dépendant connus sont (1) régulés avec une valeur-p ajustée > 0.1 (Figure 43, panneau de droite), (2) non-exprimés ou (3) non-régulés. Ces derniers possèdent au moins un autre régulateur et/ou n'ont qu'un site putatif de liaison à PhoP et aucun résultat expérimental.

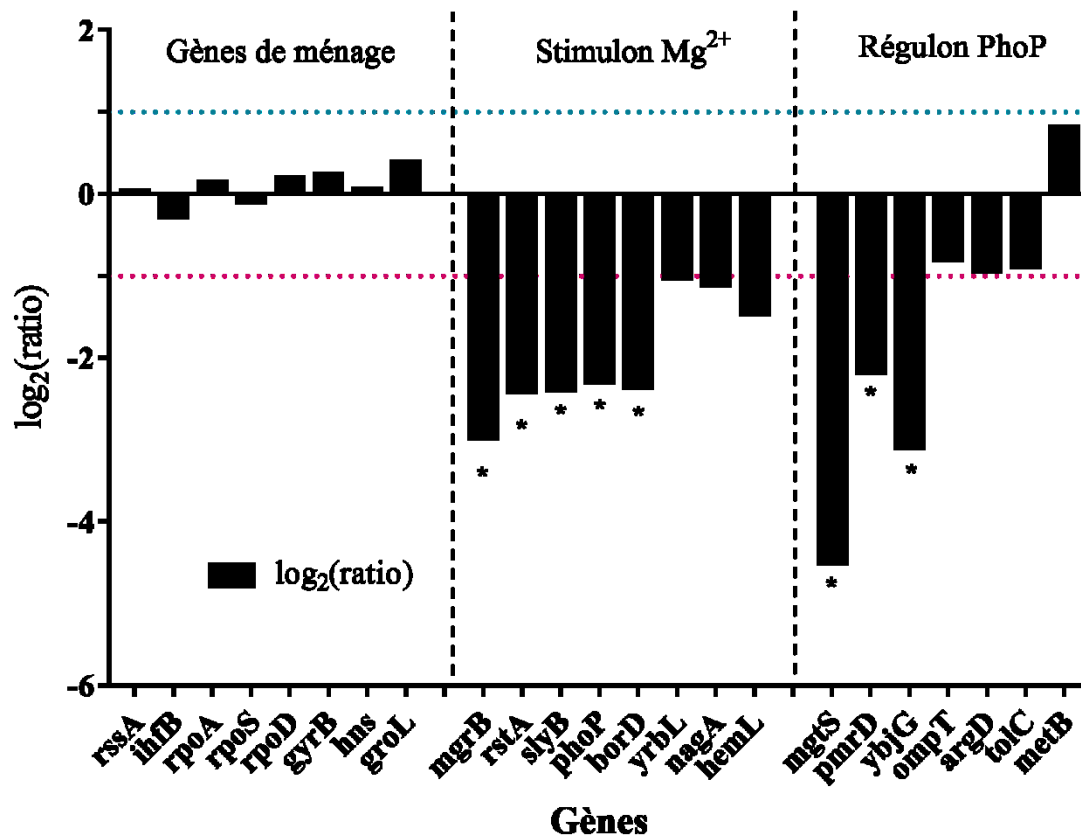


Figure 43. Ratios faible sur fort Mg^{2+} des contrôles. Les gènes de ménage ne sont pas régulés par le Mg^{2+} . Les résultats représentent le \log_2 de la moyenne et l'écart-type de deux expériences indépendantes. Ceux du stimulon Mg^{2+} et du régulon PhoP le sont. L'étoile indique une valeur-p ajustée inférieure à 0,05. La ligne pointillée bleue indique une augmentation du nombre de lecture de deux fois et la rouge une diminution de deux fois.

Dans l'ensemble, ces résultats montrent que l'ajout de Mg^{2+} a un effet ciblé sur l'expression génique du stimulon Mg^{2+} . Les données du *ribosome profiling* peuvent ainsi potentiellement servir à détecter d'autres gènes du stimulon Mg^{2+} .

5.5. IDENTIFICATION DE NOUVEAUX GÈNES DU STIMULON Mg^{2+}

L'objectif de cette partie est d'identifier de potentiels candidats au stimulon Mg^{2+} qui auraient jusqu'à présent été négligés. Pour cela, les gènes ont été classés de la valeur-p ajustée la plus basse à la plus haute. Par convention, les gènes sont considérés statistiquement significativement régulés si la valeur-p ajustée obtenue est inférieure 0,1. Le tableau 5 présente ainsi les résultats pour les gènes possédant une valeur-p ajustée $< 0,1$ à l'exception du gène *corA* qui possède une valeur-p ajustée de 0,12. Le gène *corA* a été conservé car (1) il code pour un transporteur de Mg^{2+} et (2) le gène *hemL*, qui fait partie du stimulon Mg^{2+} d'après les travaux de Minagawa *et al.* (2003), a une valeur-p ajustée de 0,15.

Tableau 5. Gènes régulés significativement par le Mg^{2+} d'après le profilage ribosomique.

	baseMean	log2FoldChange	lfcSE	padj	Régulation par
<i>yneM</i>	5460	-4.50	0.57	2.14E-12	PhoP
<i>ybjG</i>	392	-3.09	0.55	1.14E-05	PhoP
<i>mgrB</i>	201	-2.98	0.55	2.27E-05	PhoP
<i>mgtA</i>	2151	-2.38	0.46	5.52E-05	Riboswitch
<i>slyB</i>	5837	-2.38	0.46	6.36E-05	PhoP
<i>phoP</i>	834	-2.29	0.47	2.27E-04	PhoP
<i>ybeX</i>	83	2.60	0.55	4.10E-04	Inconnu
<i>rstA</i>	178	-2.40	0.54	1.56E-03	PhoP
<i>pmrD</i>	132	-2.18	0.52	3.84E-03	PhoP
<i>borD</i>	628	-2.36	0.57	4.12E-03	PhoP
<i>ynfB</i>	669	-1.82	0.45	6.87E-03	Inconnu
<i>hisD</i>	740	-1.77	0.44	7.90E-03	Inconnu
<i>rstB</i>	63	-2.28	0.58	0.011	PhoP

<i>ybjX</i>	867	-2.09	0.54	0.011	Inconnu
<i>gsiA</i>	128	1.86	0.51	0.028	Inconnu
<i>mgtL</i>	604	-1.58	0.46	0.047	Riboswitch
<i>mlaF</i>	204	-1.53	0.46	0.072	Inconnu
<i>cysD</i>	765	1.77	0.53	0.072	Inconnu
<i>ompX</i>	6015	-1.46	0.44	0.076	Inconnu
<i>hisC</i>	507	-1.40	0.43	0.086	Inconnu
<i>hisG</i>	596	-1.48	0.46	0.095	Inconnu
<i>ybeZ</i>	179	1.48	0.47	0.101	Inconnu
<i>gsiB</i>	144	1.70	0.54	0.105	Inconnu
<i>corA</i>	541	-1.29	0.42	0.116	Inconnu

Pour chaque gène sont indiqués dans l'ordre : son nom, la valeur moyenne du nombre de lectures dans les quatre échantillons, le \log_2 du ratio faible/fort Mg^{2+} , l'écart-type, la valeur-p ajustée et l'acteur de la régulation.

Parmi les 24 gènes régulés, 9 le sont par PhoP, deux par le riboswitch *mgtA*. Les 13 autres (9 unités de transcription) ne sont pas connus pour être régulés par le Mg^{2+} . L'étude suivante propose d'essayer de confirmer les résultats obtenus.

5.5.1. LE GÈNE CORA

Le gène *corA* code pour un canal de la membrane permettant le transport de plusieurs ions : Mg^{2+} , Co^{2+} , Ni^{2+} . Chez *E. coli*, il est essentiellement impliqué dans l'import et l'export de Mg^{2+} (Guskov et Eshaghi, 2012). Récemment, une étude a montré que la région 5' non traduite de l'ARNm de *corA* peut adopter deux conformations mutuellement exclusives et qu'elle est hautement conservée (Kriner et Groisman, 2015). Les auteurs montrent également que la structure adoptée par l'ARNm contrôle une terminaison de la transcription dépendante de la protéine Rho. Ces résultats suggèrent la possible présence d'un riborégulateur, mais

aucun ligand n'a été identifié à ce jour. Bien qu'aucune régulation de l'expression de *corA* n'ait été observée, les résultats du profilage ribosomique indiquent une éventuelle régulation par le Mg^{2+} , qui pourrait être le ligand du riborégulateur comme pour *mgtA*.

Afin de reproduire les résultats observés et identifier le niveau de régulation, la construction de 4 fusions différentes a été prévue :

- 1) Une fusion traductionnelle avec le promoteur endogène pour confirmer les résultats.
- 2) Une fusion transcriptionnelle avec le promoteur endogène pour mettre en évidence les cas dans lesquels la régulation se fait uniquement au niveau traductionnel.
- 3) Une fusion traductionnelle avec le promoteur constitutif *PlacUV5* pour évaluer l'importance de la région promotrice dans la régulation.
- 4) Une fusion transcriptionnelle avec le promoteur *PlacUV5* pour identifier les cas pour lesquels la régulation est faite seulement au niveau de l'ARNm.

Dans le cas de *corA*, le +1 de transcription prédit est probablement incorrect puisqu'il se trouve dans le cadre de lecture *corL*, le peptide leader de *corA*. Dans l'objectif de conserver la région promotrice complète, 700 nucléotides en amont du cadre de lecture *corA* ont été gardés pour les fusions avec le promoteur endogène. Pour les constructions avec le promoteur *PlacUV5*, le même +1 qu'identifié chez *S. enterica* a été choisi (Kriner et Groisman, 2015). Comme pour les fusions *mgtA*, uniquement les 8 premiers codons ont été fusionnés en phase avec *lacZ*. CorA est une protéine membranaire qui contient probablement un peptide signal en N-terminal. Un peptide signal étant d'une longueur habituelle de 20 à 30 résidus d'acides aminés, 8 ne devraient pas suffire à exporter la fusion de fusion CorA'-LacZ à la membrane. Malheureusement seules les fusions transcriptionnelles ont été construites avec succès. Les résultats des essais β -galactosidase sont présentés sur la Figure 44. Les activités spécifiques obtenues sont faibles en comparaison avec d'autres fusions transcriptionnelles, en particulier

pour la fusion avec le promoteur endogène. Le Mg^{2+} n'a qu'un faible effet étant donné la taille des barres d'erreur qui se chevauchent presque. Le fait que la fusion avec le promoteur constitutif semble avoir un plus grand effet que la fusion avec le promoteur endogène suggère que l'effet n'est pas observé n'est pas significatif.

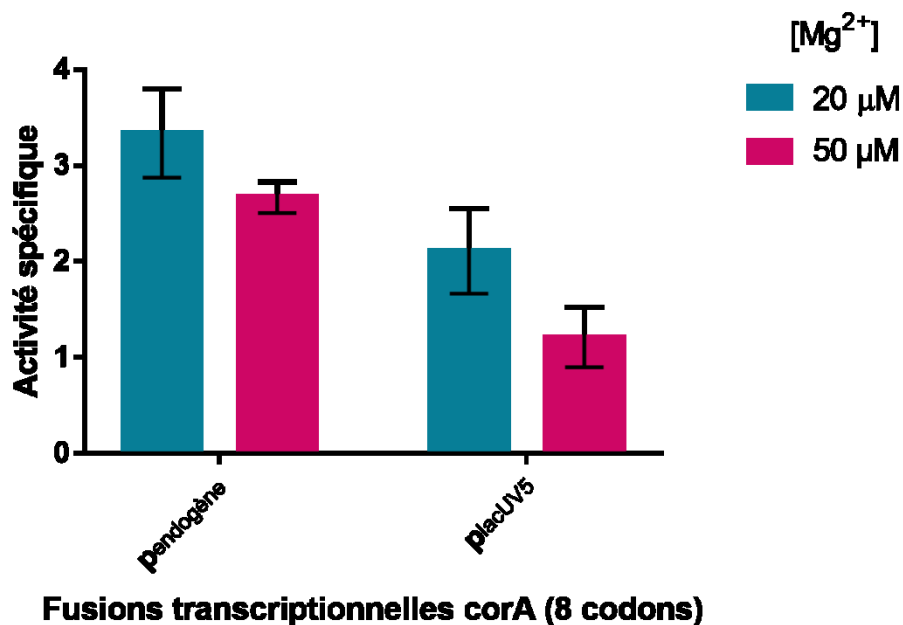


Figure 44. Effet du Mg^{2+} sur le niveau d'ARNm de *corA*. Essais β -galactosidase avec des fusions transcriptionnelles *corA* avec ses huit premiers codons sous contrôle du promoteur endogène ou de *PlacUV5*.

Plusieurs hypothèses peuvent expliquer l'absence d'effet observé au niveau de l'activité β -galactosidase :

- L'effet observé en profilage ribosomique n'est pas significatif. La valeur-p ajustée étant élevée (0,12), il est possible que la différence observée soit due à la variation normale entre les échantillons. Cette hypothèse est appuyée par l'absence de preuve dans la littérature que l'expression de *corA* pourrait être régulée par le Mg^{2+} .

- La régulation par le Mg^{2+} n'agit qu'au niveau de la traduction. Dans ce cas il est normal de ne pas observer une régulation avec une fusion transcriptionnelle. Une fusion traductionnelle, quel que soit le promoteur, devrait alors suffire à confirmer l'hypothèse.
- La régulation requiert des éléments présents dans le cadre de lecture, au-delà des 8 codons présents dans les fusions construites. Dans ce cas, un immunobuvardage de type western pourrait être plus intéressant qu'une fusion.

Il est intéressant de noter que la comparaison des profils des empreintes de ribosome dans *Genome Browser* semble indiquer que l'expression du peptide leader *corL* en amont de *corA* est également réprimée par le Mg^{2+} (Figure 45). Le peptide leader *corL* chez *E. coli* a été récemment découvert grâce à la technique de *ribosome profiling* (Oh et al., 2011) et son implication dans le contrôle de l'expression de *corA* a été établie chez *S. enterica* (Kriner et Groisman, 2015). Afin d'évaluer son rôle chez *E. coli*, l'outil développé au chapitre 4 pour étudier le rôle du peptide leader *mgtL* a été utilisé pour *corL*. Ainsi une fusion transcriptionnelle *thiM-corL-corA* a été construite et des essais β -galactosidase ont été réalisés (Figure 46). À l'inverse de *mgtL*, la traduction de *corL* semble réprimer l'expression de *corA* puisque l'activité est plus élevée avec TPP. Cette anticorrélation entre *corL* et *corA* ne s'observe pas sur les profils de la Figure 45. Il semblerait donc que l'effet du Mg^{2+} observé sur l'expression de *corA* soit dû à une régulation au niveau de l'ARNm.

Malgré des résultats peu prometteurs pour *corA*, d'autres expériences sont nécessaires pour confirmer ou infirmer l'effet du Mg^{2+} sur son expression. Le lien observé entre l'expression de *corA* et de son peptide leader semble indiquer un mécanisme de régulation par le peptide et confirme la pertinence de l'outil développé au chapitre 4 pour évaluer l'impact de la traduction d'un peptide leader sur le gène en aval.

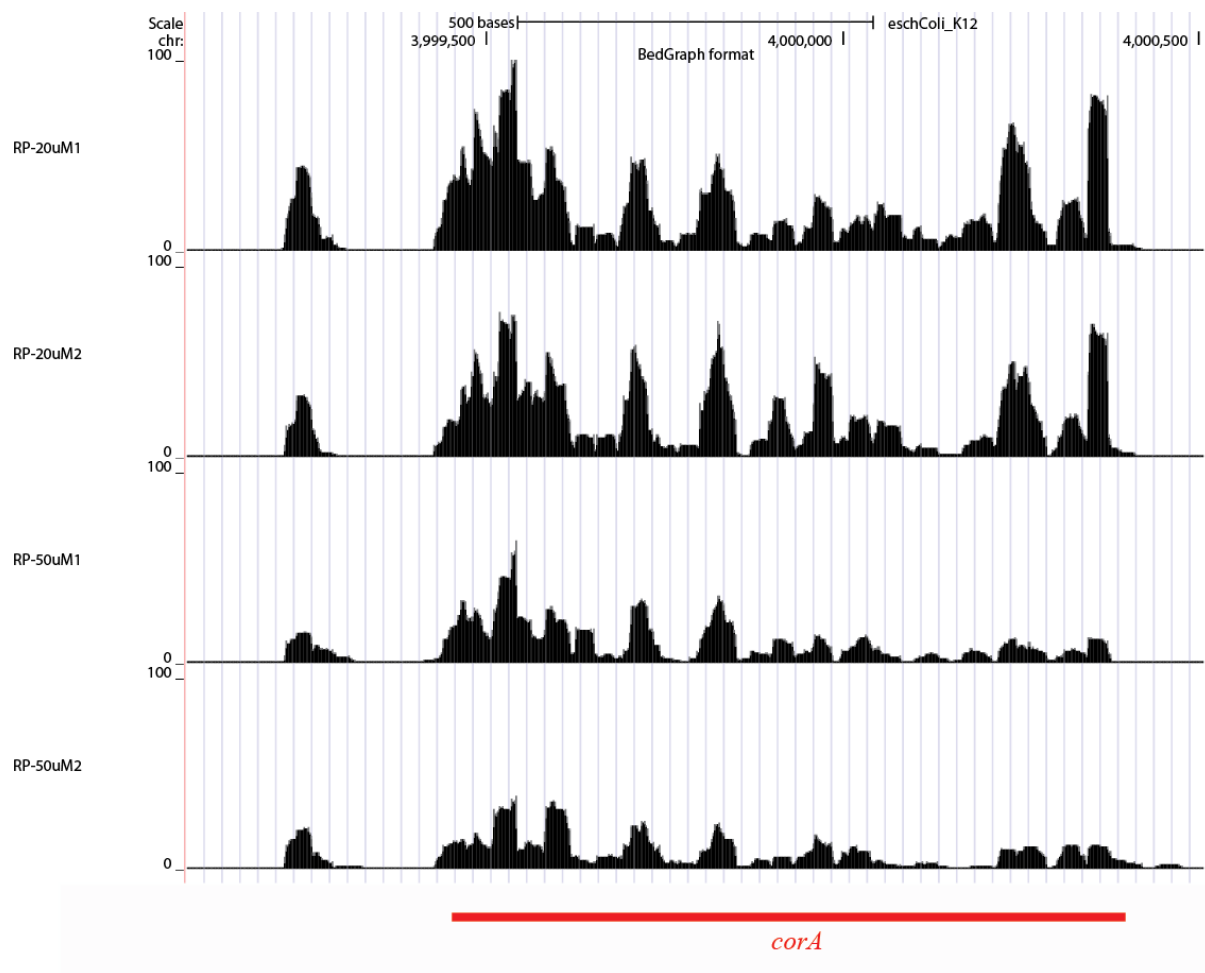
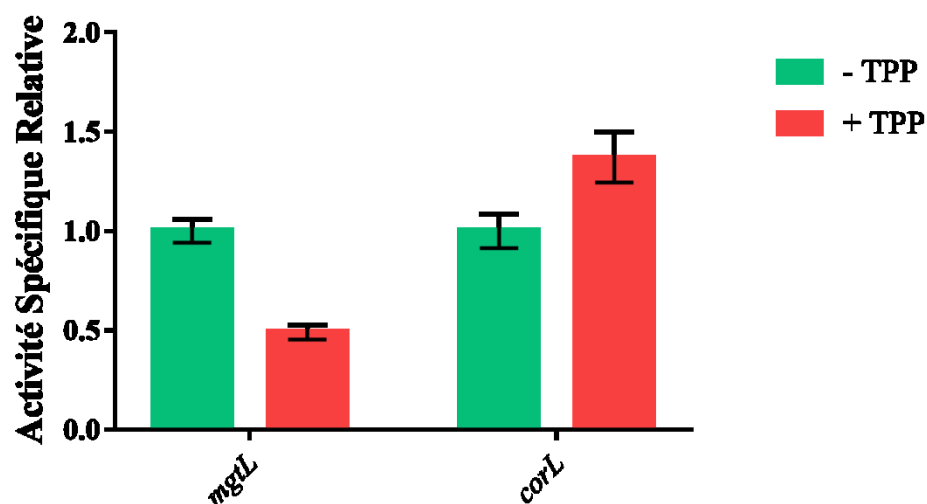


Figure 45. Résultats du profilage ribosomique pour *corA*. Nombre de lecture en fonction de la position, pour les quatre échantillons.



Fusions hybrides du riborégulateur *thiM* avec le peptide indiqué

Figure 46. Effet de la traduction de *corL* sur l'expression de *corA*. Essais β -galactosidase réalisés avec des fusions hybrides *thiM-mgtL-mgtA'-lacZ* et *thiM-corL-corA'-lacZ*. Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

5.5.2. LE GÈNE YBEX

Le gène *ybeX* est le plus significatif parmi les candidats retenus. De plus c'est un des rares gènes à être activé par le Mg^{2+} . Peu d'informations sont disponibles sur *ybeX* mais la protéine présente une grande homologie avec la protéine CorC qui permet l'efflux de Mg^{2+} et de Co^{2+} (Gibson et al., 1991). Ainsi à l'inverse du transporteur MgtA qui importe le Mg^{2+} et qui est produit lorsque la concentration intracellulaire en Mg^{2+} est faible, une forte concentration activerait la production du transporteur YbeX (ou CorC) qui exporterait l'excédent en Mg^{2+} (Figure 47).

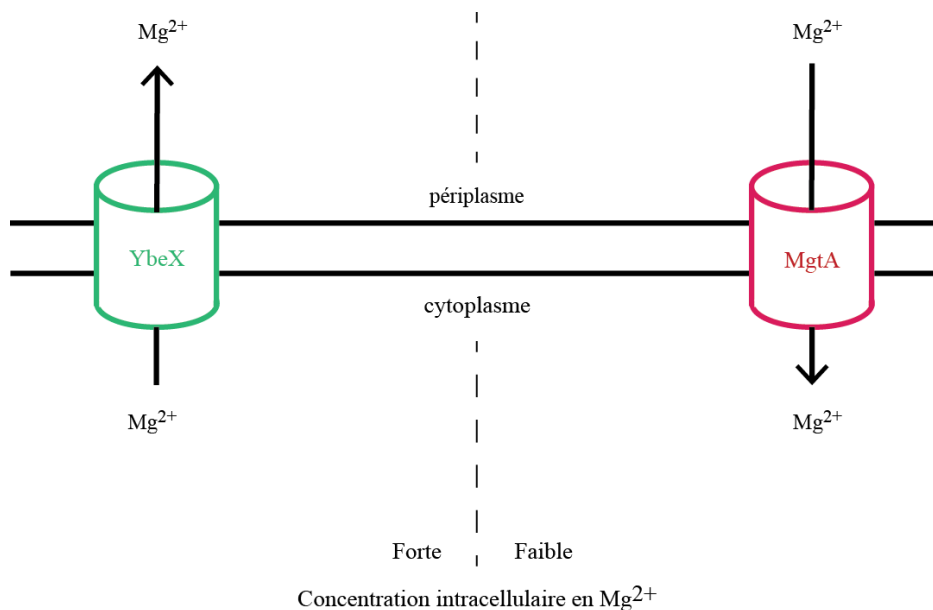


Figure 47. Comparaison des transporteurs MgtA et YbeX. MgtA permet l'import de Mg^{2+} dans le cytoplasme en cas de carence, YbeX est prédit pour l'exporter en cas de surplus.

Le gène *ybeX* pourrait faire partie de plusieurs unités de transcription distinctes (Figure 48). Le gène en amont *ybeZ* étant également activé par le Mg^{2+} selon le *ribosome profiling* (Tableau 6), il est possible que *ybeX* fasse partie de l'opéron *ybeZYX*. Comme *ybeY* ne fait pas partie de la liste des gènes les plus régulés et que *ybeZ*, premier gène de l'opéron, est moins régulé que *ybeX* (avec un ratio en \log_2 de 1,48 et 2,60 respectivement), il est très envisageable que la régulation de *ybeX* soit indépendante du reste de l'opéron. Un promoteur σ^{32} a été identifié en amont de *ybeZ*. Le facteur σ^{32} est généralement associé à un choc thermique ou à un choc osmotique. Aucun autre gène dépendant du facteur σ^{32} n'a été identifié parmi les gènes régulés par le Mg^{2+} , la possibilité d'un effet global σ^{32} peut donc être écartée.

Gene Local Context (not to scale -- see Genome Browser for correct scale) ?

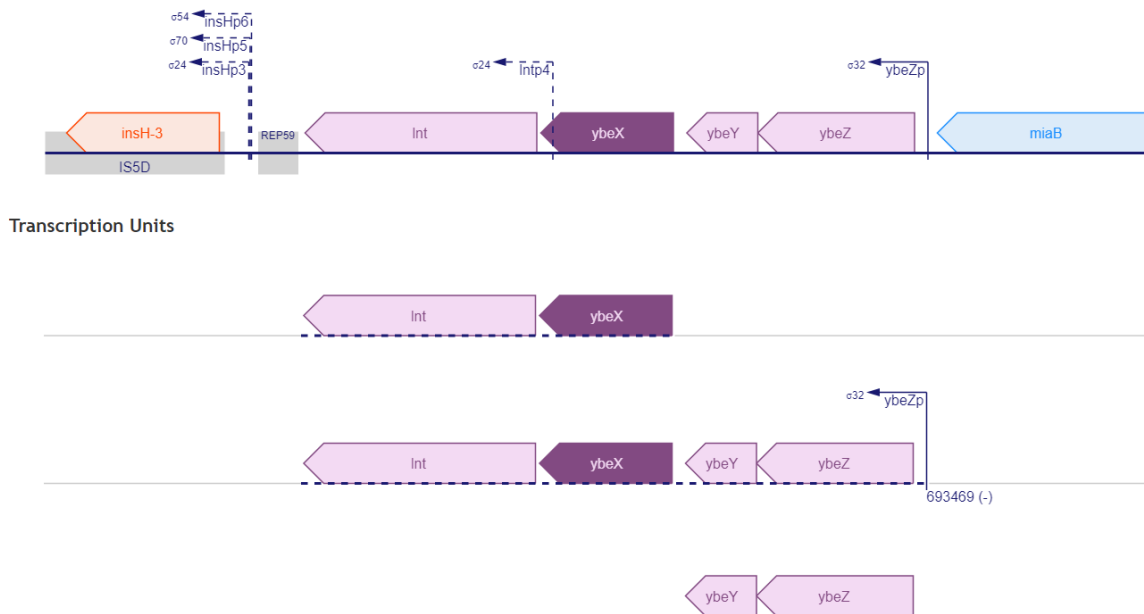


Figure 48. Contexte génomique du gène *ybeX*. Le seul promoteur identifié permet la transcription de *ybeX* en opéron avec *ybeZ* et *ybeY*. Aucun terminateur n'a été identifié entre *ybeX* et *Int*, suggérant que les deux gènes font partie de la même unité de transcription. Un transcrit *ybeX-Int* pourrait exister. Données du site Ecocyc.

Pour confirmer les résultats de *ribosome profiling* et identifier le niveau de régulation, les fusions prévues incluent ce promoteur et la possibilité d'un promoteur proximal pour transcrire *ybeX* indépendamment de *ybeZY*. YbeX étant prédite pour être une protéine membranaire, seuls les 3 premiers codons ont été conservés pour la fusion avec *lacZ*. Seule la fusion transcriptionnelle avec le promoteur proximal a été obtenue. Les résultats des essais de β -galactosidase montrent que ce promoteur putatif existe probablement puisqu'une activité a été détectée (Figure 49). Cette activité est cependant très faible. Une légère répression est observée au lieu de l'activation attendue. Ces résultats suggèrent que la régulation pourrait se faire à partir d'un promoteur distal ou que la régulation se fait au niveau traductionnel ou qu'une plus grande région du cadre de lecture est nécessaire pour la régulation. D'autres fusions devront être construites pour déterminer la ou les raisons pour lesquelles les résultats du *ribosome profiling* n'ont pas été confirmés.

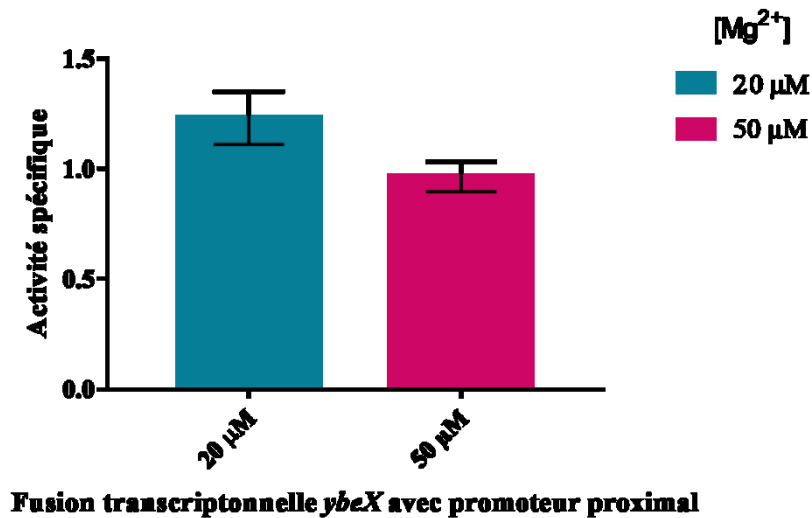


Figure 49. Effet du Mg^{2+} sur le niveau d'ARNm de *ybeX*. Essais β -galactosidase réalisés avec la fusion transcriptionnelle *ybeX* contenant les quatre premiers codons et sous le contrôle du promoteur proximal putatif. Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

5.5.3. LE GÈNE *YNFB*

Le gène *ynfB* est réprimé 3,53 fois d'après les données de profilage ribosomique. La fonction de YnfB n'est pas connue, mais des études ont montré qu'elle est localisée à la membrane externe (Lopez-Campistrous et al., 2005) et qu'elle pourrait jouer un rôle dans la biogenèse de la membrane cellulaire (Paradis-Bleau et al., 2014).

Une fusion traductionnelle contenant les 5 premiers codons de *ynfB* et sous le contrôle du promoteur endogène a été construite. Les résultats obtenus en essais β -galactosidase confirment les données de séquençage avec une diminution de l'activité en présence de plus de Mg^{2+} (Figure 50). Comme aucune fusion sous le contrôle du promoteur *PlacUV5* n'a été obtenue, la possibilité d'une régulation au niveau du promoteur a été envisagée en cherchant des boîtes PhoP. Le consensus est une répétition de (T/G)GTTTA séparées par 5 paires de base, environ 25 pb en amont du +1 de transcription. Le +1 est localisé 99 nucléotides avant

l'AUG. Seule une répétition a été trouvée au début de la séquence transcrite. Le promoteur de *ynfB* ne semble donc pas être directement dépendant de PhoP. PhoP contrôle d'autres systèmes à deux composants tels que RstA/RstB (régulation directe) ou PmrA/PmrB (régulation indirecte via PmrD). Il est donc possible que l'expression de *ynfB* soit régulée par le Mg^{2+} via la régulation d'un facteur de transcription lui-même sous contrôle de PhoP. C'est la première fois que la régulation de *ynfB* par le Mg^{2+} a été reportée bien que plusieurs études transcriptomiques ont été réalisées en faible et fort Mg^{2+} . Il est possible que les conditions de l'expérience soient responsables : les autres études utilisent de plus grandes concentrations en Mg^{2+} et étudient généralement la phase exponentielle de croissance. Il est également possible que la régulation de *ynfB* soit uniquement traductionnelle, ce qui expliquerait qu'elle n'est pas été détectée avec des études transcriptomiques. La région 5' non traduite ne semble pas être conservée, la présence d'un riborégulateur Mg^{2+} est donc peu probable.

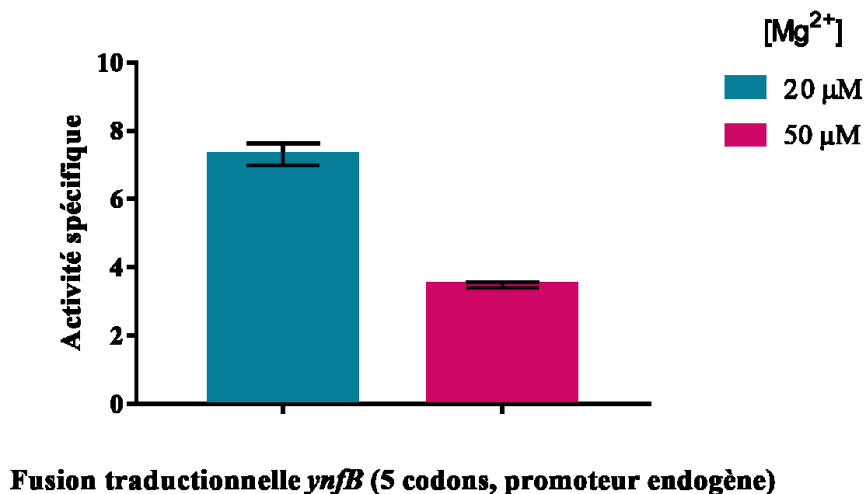


Figure 50. Effet du Mg^{2+} sur l'expression de *ynfB*. Essais β -galactosidase réalisés avec une fusion traductionnelle de *ynfB* contenant les cinq premiers codons et sous contrôle du promoteur endogène. Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

5.5.4. L'OPÉRON HISLGDC

Les gènes *hisG*, *hisD* et *hisC* sont réprimés d'un facteur 2,79, 3,41 et 2,64 respectivement. Ces trois gènes font partie du même opéron qui contient en plus le peptide leader *hisL*. Le pic correspondant à *hisL* est très faible en comparaison avec le reste de l'opéron (Figure 51). La traduction complète du peptide est responsable de l'atténuation de la transcription (Alifano et al., 1996). Les gènes en aval étant traduits et donc transcrits, il n'est donc pas étonnant de ne pas observer une forte traduction du peptide en faible Mg^{2+} . Il n'est pas possible d'évaluer le rôle du peptide *hisL* dans la régulation de l'expression de l'opéron avec seulement les données de profilage ribosomique. En effet, des données de séquençage permettraient de déterminer si la plus faible densité d'empreintes de ribosome est due à une atténuation de la transcription ou à une inhibition de la traduction.

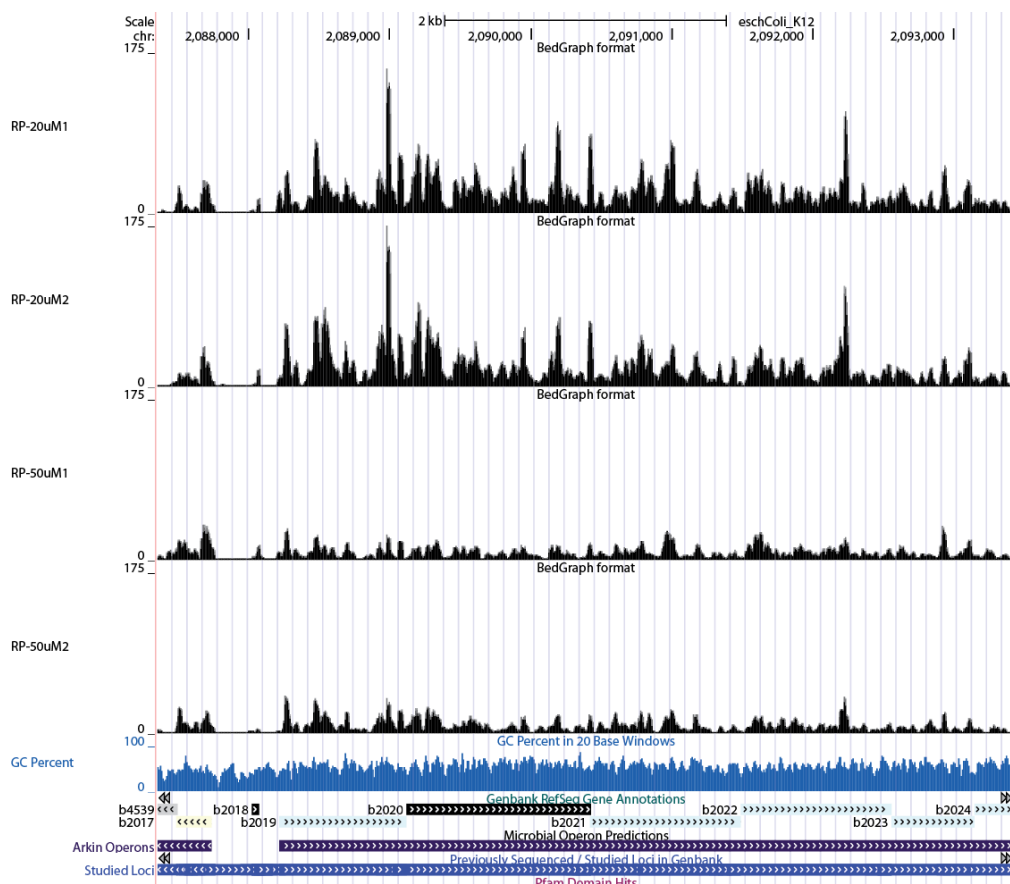


Figure 51. Résultats du profilage ribosomique pour l'opéron *hisLGDC*. Le nombre de lecture est représenté en fonction de la position pour les quatre échantillons.

5.5.5. LE GÈNE YBJX

Le gène *ybjX* est réprimé 4,26 fois d'après les données de *ribosome profiling*. Sa fonction est inconnue, mais son homologue *somA* chez *S. enterica* est important pour l'infection (Detweiler et al., 2003). L'expression de *somA* est dépendante de PhoP et des travaux suggèrent que le gène *ybjX* pourrait l'être aussi chez *E. coli* (Zwir et al., 2005). La régulation observée se ferait donc au niveau du promoteur de manière dépendante de PhoP. Le +1 de transcription identifié (Mendoza-Vargas et al., 2009) se situe après l'AUG annoté de *ybjX* et de manière intéressante, le profil de profilage ribosomique ne correspond pas à l'AUG annoté

(Figure 52). En effet, un pic centré sur l'AUG ou la région Shine-Dalgarno est généralement observé et dans le cas de *ybjX* le premier pic est dans le cadre de lecture. Les résultats de profilage ribosomique confirme ainsi que le +1 de transcription est bien à la position indiquée et que c'est l'annotation du cadre de lecture qui est fausse. Le transcrit *ybjX* contiendrait donc une région 5' non traduite très courte et ne contenant pas de séquence Shine-Dalgarno.

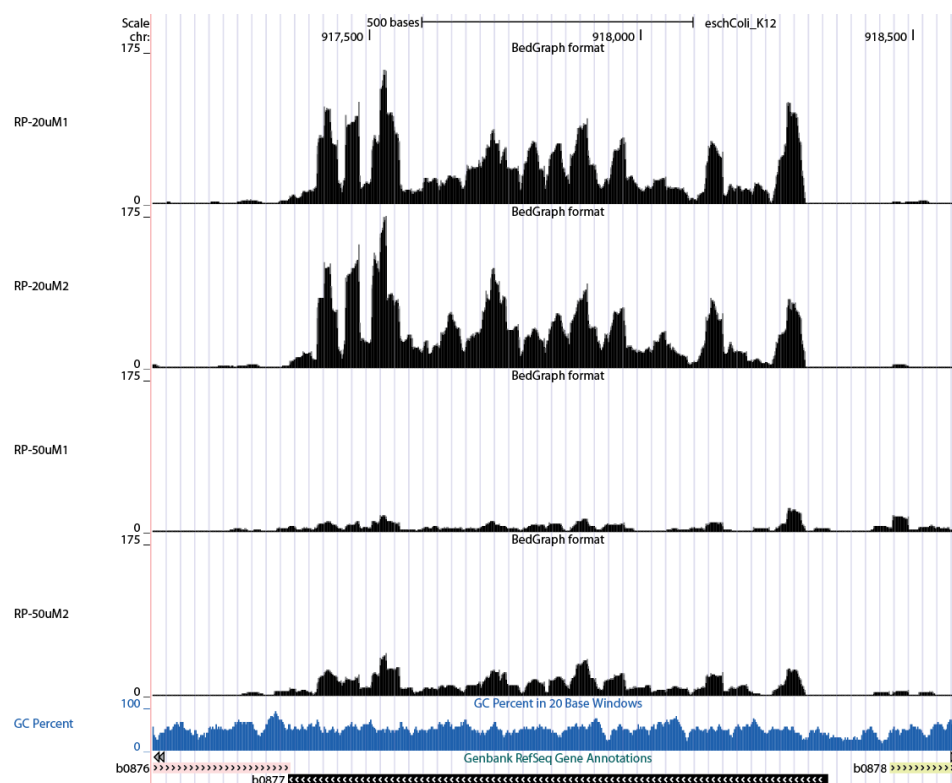


Figure 52. Résultats du profilage ribosomique pour le gène *ybjX*. Le nombre de lecture est représenté en fonction de la position pour les quatre échantillons.

5.5.6. L'OPÉRON GSIAB

Les gènes *gsiA* et *gsiB* sont activé 3,63 et 3,5 fois respectivement. Ils font partie de l'opéron *iaaA-gsiABCD*. Le gène *iaaA* est quant à lui activé 2,24 fois avec une valeur-p ajustée de 0,58.

Les profils suggèrent que *iaaA* est plus activé dans le premier réplikat et peu dans le 2^{ème}, ce qui explique la forte valeur-p ajustée (Figure 53). Les gènes *gsiC* et *gsiD* ne sont pas assez exprimés pour observer une régulation. Les gènes *gsi* codent pour des sous-unités d'un transporteur de glutathionne de type ABC. La sous-unité GsiA a une activité ATPase et GsiB est une sous-unité périplasmique.

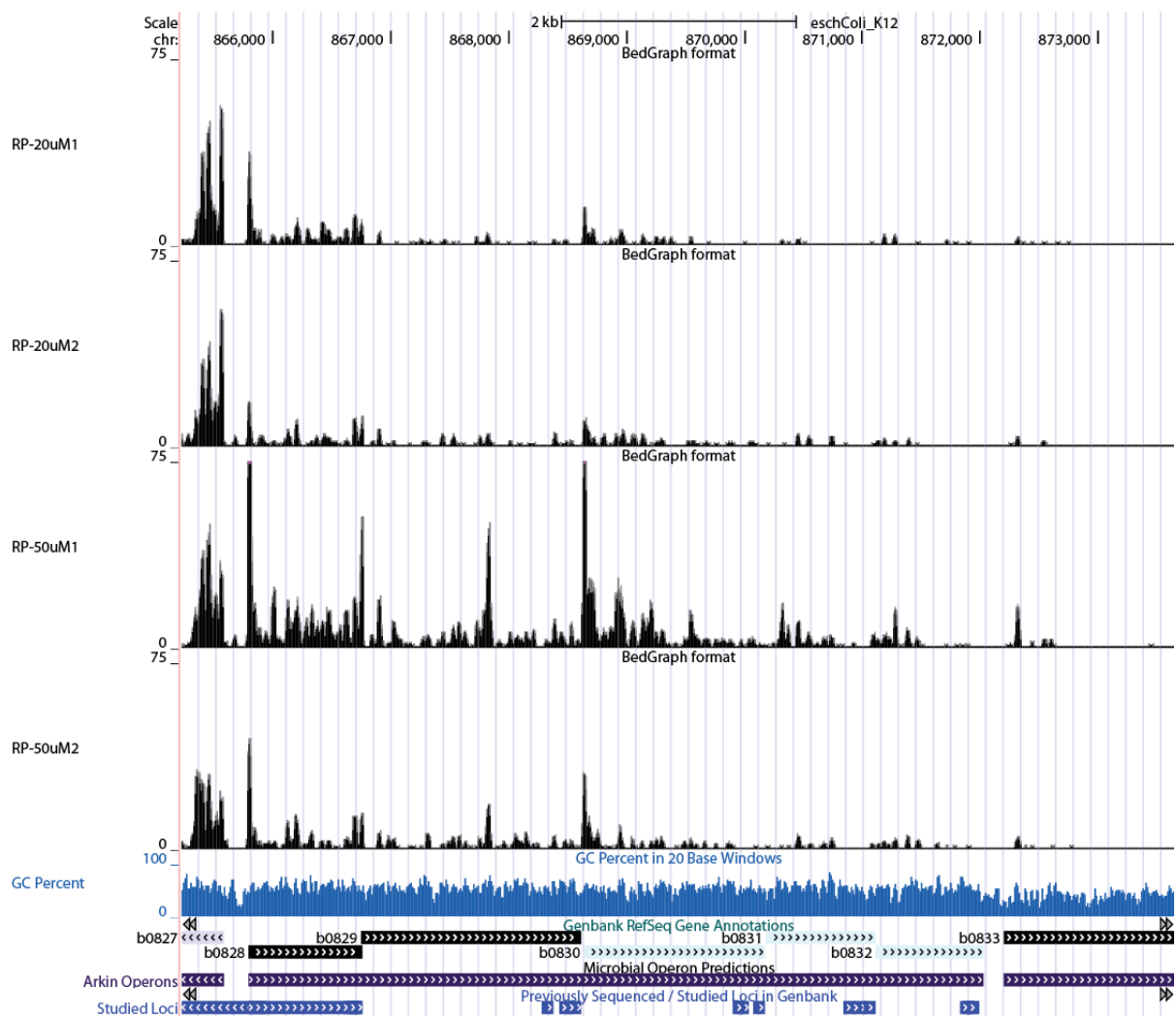


Figure 53. Résultats du profilage ribosomique pour l'opéron *gsiAB*. Le nombre de lecture est représenté en fonction de la position pour les quatre échantillons.

5.5.7. LE GÈNE *MLAF*

Le gène *miaF*, réprimé 2,88 fois, est le premier de l'opéron *miaFEDCB*. Les gènes *miaE* et *miaD* sont également réprimés, respectivement 2,18 et 2,51 fois, mais les valeurs-p ajustées sont supérieures à 0,1. Tous les gènes de l'opéron sont responsables du transport de phospholipides à travers la membrane (Ekiert et al., 2017). Les profils montrent que les gènes de l'opéron sont exprimés différemment : *miaF* et *miaE* ont des profils similaires, *miaD* est un peu moins exprimé, *miaC* est plus fortement exprimé et *miaB* ne semble que très peu exprimé (Figure 54). Il semble donc qu'il y ait une forte régulation au niveau de cet opéron. Il est ainsi possible que *miaC* ait son propre promoteur et qu'il y ait un terminateur de transcription entre *miaC* et *miaB*. Le gène *ibaG* en amont de l'opéron a un profil très semblable dans les quatre échantillons, confirmant que l'effet observé est spécifique à l'opéron *miaFEDCB*.

5.5.8. LE GÈNE *CYSD*

Le gène *cysD* est activé 3,41 fois. Il est en opéron avec *cysN* et *cysC*, activés respectivement 2,69 et 3,39 fois, mais avec des valeurs-p plus élevées que 0,1 (0,35 et 0,18). Les profils de *ribosome profiling* suggèrent que l'opéron est principalement activé dans le premier réplicat et peu ou pas dans le deuxième (Figure 55). Ce résultat pourrait être un artéfact ou être représentatif d'une variation des conditions le second jour. Des fusions sont ici nécessaires pour confirmer l'effet obtenu.

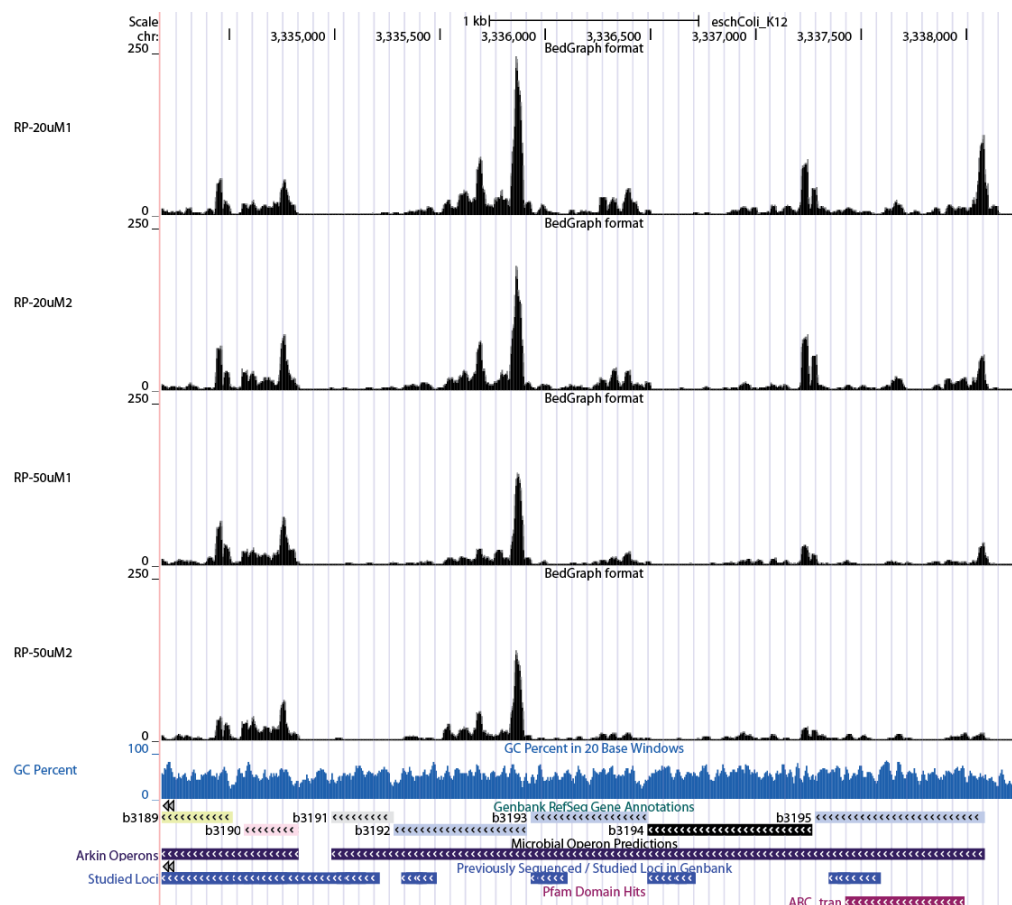


Figure 54. Résultats du profilage ribosomique pour l'opéron *malFEDCB*. Le nombre de lecture est représenté en fonction de la position pour les quatre échantillons.

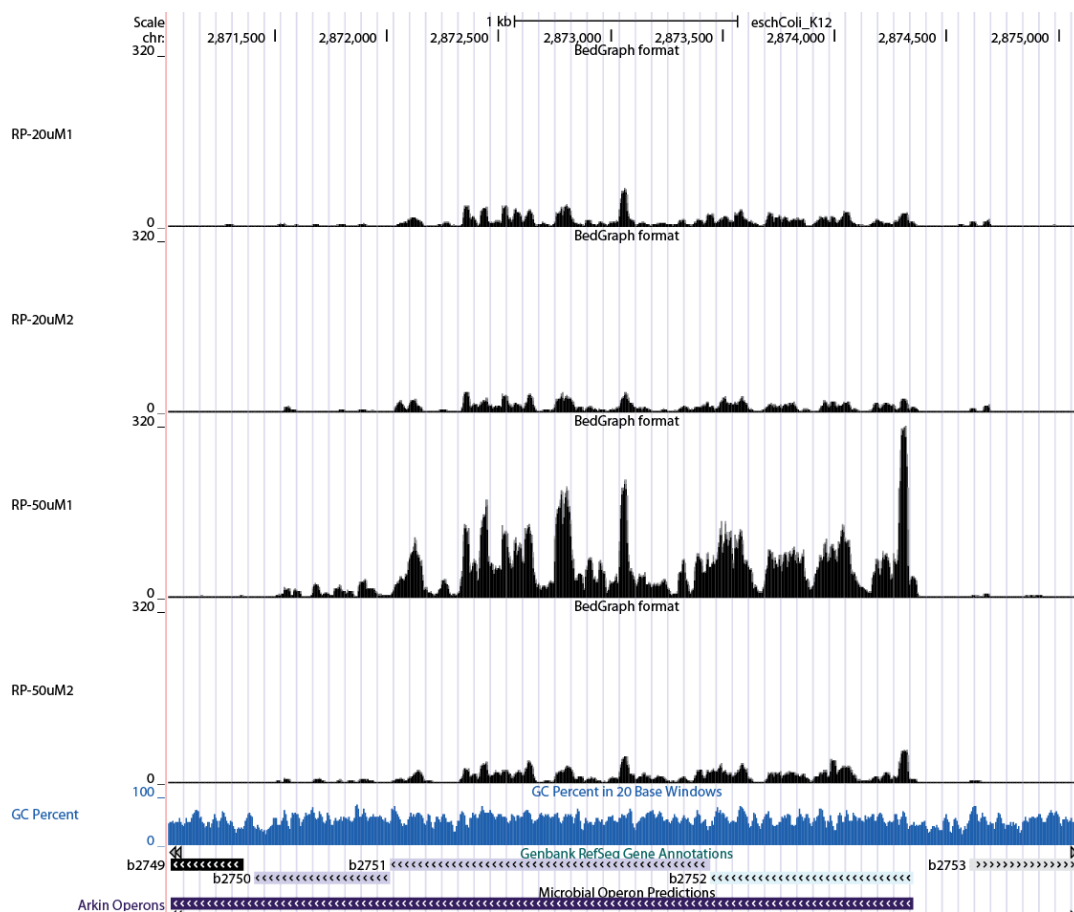


Figure 55. Résultats du profilage ribosomique pour l'opéron *cysDNC*. Le nombre de lecture est représenté en fonction de la position pour les quatre échantillons.

5.5.9. LE GÈNE *OMPX*

Le gène *ompX* est réprimé 2,75 fois. *OmpX* est une petite protéine localisée à la membrane externe comme son nom l'indique (*Outer membrane protein*). C'est une porine impliquée dans l'export de YebF dont la fonction est inconnue (Prehna et al., 2012). Par ailleurs, la surexpression de *ompX* augmente l'activité du facteur de transcription σ^E important notamment lors d'un stress membranaire (Mecsas et al., 1993). *ompX* semble ainsi jouer un rôle crucial dans l'intégrité de la membrane.

Une fusion traductionnelle contenant le(s) promoteur(s) endogène(s) (deux promoteurs dans la région amplifiée) a été construite. Les résultats des essais β -galactosidase confirment les données de profilage ribosomique avec une répression de l'expression de *ompX* en fort Mg^{2+} (Figure 56).

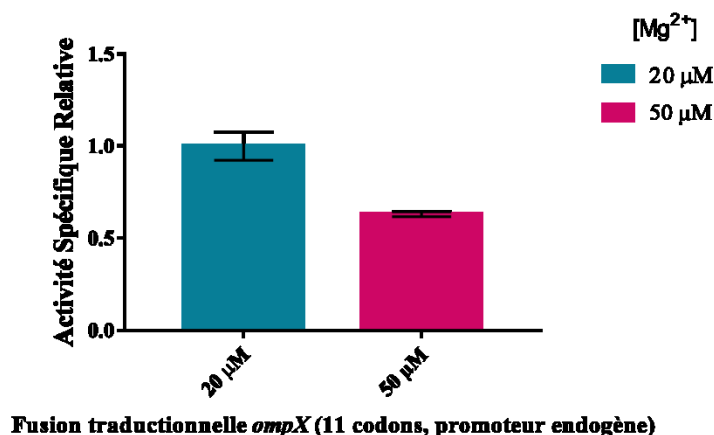


Figure 56. Effet du Mg^{2+} sur l'expression de *ompX*. Essais β -galactosidase réalisés avec une fusion traductionnelle *ompX* contenant les onze premiers codons et sous le contrôle du promoteur endogène. Les résultats représentent la moyenne et l'écart-type d'au moins deux expériences indépendantes.

Les autres fusions n'ont pas pu être obtenues, mais des évidences suggèrent que la régulation pourrait s'effectuer au niveau du promoteur de manière dépendante de PhoP. En effet, une étude visant à identifier les promoteurs régulés directement par PhoP a identifié une région qui pourrait lier PhoP (Zwir et al., 2005).

CHAPITRE 6

DISCUSSION ET CONCLUSION GÉNÉRALE

6.1. CONCLUSION GÉNÉRALE

Les études présentées dans cette thèse avaient pour objectifs de mieux comprendre les mécanismes de régulation employés par le riborégulateur *mgtA* et d'identifier d'autres gènes régulés par le Mg^{2+} , de manière indépendante du régulateur PhoP. Les résultats obtenus dans le Chapitre 3 montrent que le riborégulateur régule l'expression de *mgtA* au niveau de l'ARNm et plus précisément de l'élongation de la transcription. Le Chapitre 4 apporte de nouvelles informations à propos du rôle du peptide leader *mgtL* et le place au cœur du mécanisme de régulation du riborégulateur. Enfin l'analyse de données de *ribosome profiling* présentée Chapitre 5 a permis l'identification de gènes régulés par le Mg^{2+} qui ne le sont potentiellement pas de manière dépendante de PhoP.

Les données obtenues aux Chapitres 3 et 4 ont permis d'établir un modèle du mécanisme de régulation de *mgtA* au niveau post-transcriptionnel et seront l'objet d'un article scientifique en cours de rédaction. Puisque les résultats préliminaires ont montré que le riborégulateur de *E. coli* était effectivement fonctionnel, c'est ce dernier et non celui de *S. enterica* qui a fait l'objet de l'étude de cette thèse. Certains résultats obtenus concordent avec ceux publiés chez *S. enterica* tels que l'implication d'un mécanisme transcriptionnel. Cependant, d'autres résultats montrent des divergences entre les deux organismes, comme l'implication de la protéine Rho, et d'autres encore ont mené à revoir l'interprétation des résultats obtenus chez *S. enterica*, comme la régulation de la traduction de *mgtL*. Le modèle du mécanisme de régulation établi ainsi sera discuté dans les paragraphes suivants.

Le Chapitre 5 présente une analyse statistique des données de séquençage obtenues sur les fragments d'ARN protégés par les ribosomes. Cette technique permet une meilleure appréhension de l'expression génique que le séquençage des transcrits. Malgré le manque de réplicats et des valeurs de valeur-p élevées, l'expression différentielle révèle que le Mg^{2+} régule plus de gènes que ce qui a été publié à ce jour.

6.2. Mg^{2+} ET PHASE DE CROISSANCE

Lors de la détermination des conditions d'observation de la régulation de l'expression de *mgtA* par le riborégulateur, il a été noté que la phase de croissance était très importante. En effet, peu de régulation par le Mg^{2+} n'est observée en phase exponentielle. De plus, cette régulation augmente dès la fin de la phase exponentielle et continue d'augmenter lorsque la phase stationnaire est atteinte. Les courbes Activité Spécifique = $f(DO_{600})$ montrent que l'expression de *mgtA* est réprimée en phase exponentielle quelle que soit la concentration de Mg^{2+} (Figure 15, panneaux B et C). En fin de phase exponentielle, l'expression de *mgtA* n'augmente que pour les concentrations faibles en Mg^{2+} . Ce résultat suggère que l'expression de *mgtA* nécessite une activation ou une dérégulation. Deux hypothèses ont été avancées pour expliquer l'influence de la phase de croissance. La première hypothèse est qu'un facteur protéique exprimé de manière dépendante de la phase de croissance est responsable de la régulation. Par exemple, les gènes *bmrBC* chez *B. subtilis* sont exprimés spécifiquement en fin de phase exponentielle de croissance grâce au facteur de transcription AbrB qui répriment leur expression au cours de la phase exponentielle en agissant au niveau de leur promoteur (Reilman et al., 2014). L'autre hypothèse est que la régulation ne se fait qu'en fin de phase exponentielle parce que la concentration en Mg^{2+} intracellulaire libre diminue lors de cette même phase. En effet, une étude a montré que la concentration intracellulaire en Mg^{2+} libre diminue lors du passage de la phase exponentielle à la phase stationnaire et que cette diminution est plus importante en faible Mg^{2+} (10 μ M, diminution d'un facteur 2) qu'en fort Mg^{2+} (1mM, diminution d'un facteur 1,3) (Lusk et al., 1968). Ainsi, l'expression de *mgtA*

augmenterait uniquement en faible Mg^{2+} et en fin de phase exponentielle parce qu'à cette concentration et à cette phase de croissance, la concentration intracellulaire en Mg^{2+} libre diminue, modulant alors la structure du riborégulateur. Le fait que la fusion hybride *thiM-mgtL-mgtA* permette de réguler l'expression de *mgtA* par la TPP en phase exponentielle suggère que la dépendance de la phase de croissance ne vient pas d'un facteur exprimé de manière dépendante de la phase de croissance, mais plutôt du signal lui-même, soit que la disponibilité en Mg^{2+} est elle-même dépendante de la phase de croissance, ce qui n'est pas le cas de la TPP.

6.3. RÔLE DU PEPTIDE LEADER

Le riborégulateur *mgtA* semble bien correspondre au modèle de riborégulateur dans lequel le gène régulé code pour un gène de synthèse ou de transport du ligand reconnu puisque le gène *mgtA* code pour un transporteur de Mg^{2+} , ligand du riborégulateur. Cependant, le riborégulateur *mgtA* présente une différence majeure qui est l'implication d'un peptide leader pour réguler l'expression du gène en aval. Au contraire, les autres riborégulateurs chez *E. coli* ne requièrent que leur structure et leur ligand pour réguler. Pourquoi la traduction d'un cadre de lecture est-elle nécessaire à l'expression de *mgtA*? Les peptides leaders sont généralement présents dans les régions 5' de gènes ayant un lien avec la traduction : biosynthèse d'un acide aminé, résistance à un antibiotique ciblant le ribosome, *etc.* La protéine MgtA a-t-elle un lien avec la traduction? Bien que cette protéine ait été identifiée comme transporteur de Mg^{2+} , plusieurs questions sur son rôle précis demeurent. En effet, *E. coli* possédant un autre transporteur de Mg^{2+} , CorA, la nécessité d'un deuxième transporteur suggère un rôle différent pour chacune de ces protéines. CorA est considéré comme le principal transporteur en raison de sa conservation parmi les eubactéries et de sa plus forte capacité (Smith et Maguire, 1998). De plus, le transporteur MgtA semble permettre l'influx de Mg^{2+} en utilisant son activité ATPase dans le sens et non à l'encontre du gradient électrochimique. Ceci suggère qu'un substrat secondaire pourrait être co- ou contre-transporté par MgtA ou que l'import de Mg^{2+} et

l'énergie de l'ATP permettrait le transport d'un autre substrat. Le rôle du peptide leader pourrait ainsi être en lien avec cette fonction secondaire de la protéine MgtA. Cependant, aucune molécule n'ayant été identifiée comme substrat de MgtA, cette hypothèse reste peu plausible. Le lien entre MgtA et la traduction pourrait être plus général, le Mg^{2+} étant important pour le processus de traduction et notamment pour le maintien de l'intégrité des ribosomes (Gesteland, 1966; McCarthy, 1962). Une baisse de l'efficacité de la traduction pourrait alors être détectée par *mgtL*, le *calage* du ribosome induisant l'expression de MgtA pour importer plus de Mg^{2+} dans la cellule et permettre de pallier la baisse de performance de la traduction.

La séquence du peptide leader a été démontrée comme étant importante pour la régulation. Le rôle de cette séquence peut être de détecter tout changement d'efficacité de la traduction en induisant le *calage* du ribosome et donc l'expression de *mgtA*. La présence de plusieurs codons proline a mené une équipe de recherche à poser l'hypothèse que le peptide leader permettrait en fait de détecter spécifiquement une baisse de la quantité de prolines, soit la région 5' de *mgtA* utilisent deux signaux et deux mécanismes pour réguler *mgtA* (Park et al., 2010). Cette équipe a ainsi examiné le lien entre le niveau de proline et MgtA. Ils ont ainsi montré que la protéine MgtA n'est pas un transporteur de proline et qu'elle n'est pas nécessaire à la biosynthèse de proline. Ils ont avancé l'idée que la proline peut être un osmoprotecteur pour la cellule et que le Mg^{2+} peut être nécessaire pour protéger les membranes et les ribosomes en cas de choc osmotique. La proline étant moins disponible pour la traduction, le *calage* du ribosome sur *mgtL* permettrait l'expression de *mgtA* pour augmenter la concentration en Mg^{2+} . La régulation de *mgtA* par le niveau de proline a cependant été discutée (Zhao et al., 2011).

6.4. COMPARAISON AVEC *S. ENTERICA*

E. coli et *S. enterica* sont phylogénétiquement proches puisqu'elles font partie de la famille des *Enterobacteriaceae* et possèdent 68% de gènes homologues (McClelland, 2000). Le gène *mgtA* est ainsi présent dans les deux organismes et leur expression est régulée par le Mg^{2+} via leur région 5'. La structure du riborégulateur est d'ailleurs conservée et la séquence du 5' possède 69% d'homologie. Le cadre de lecture de *mgtL* est également présent dans les deux séquences et la séquence peptidique est similaire à 70%. Malgré ces caractéristiques partagées le mécanisme de régulation semble différent puisque le riborégulateur chez *E. coli* n'est pas dépendant de Rho alors que celui de *S. enterica* l'est. Ce résultat inattendu a été étudié avec soin, en utilisant différentes techniques (essais d'immunobuvardage de type *northern*, β -galactosidase, immunoprécipitation de la chromatine), un inhibiteur (la bicyclomycine) ou un mutant (de Rho), dans deux laboratoires (les laboratoires du Pr Daniel Lafontaine à l'Université de Sherbrooke et du Pr Joseph Wade du Wardworth Center à Albany). Des études récentes ont également confirmé ces résultats, avec une combinaison de séquençage des 3' et séquençage des ARN totaux en présence et en absence de bicyclomycine, montrant qu'un site de terminaison de la transcription entre *mgtL* et *mgtA* n'était pas dépendant de Rho (Dar et Sorek, 2018; Peters et al., 2012). Cette divergence dans les mécanismes peut venir de la séquence du riborégulateur ou de la protéine Rho. Le gène *rho* est présent dans plus de 90% des génomes bactériens séquencés et est essentiel pour la majorité des bactéries à Gram négatif telles que *E. coli* et *S. enterica* (D'Heygère et al., 2013). Les séquences protéiques de Rho étant à 99% identiques, il est peu probable que la différence de protéine Rho soit à l'origine de la différence de mécanisme. Il est plus probable que la variation soit due à la séquence du riborégulateur. Il n'existe pas cependant de site rut consensus pour le riborégulateur *mgtA* chez *S. enterica*, il est donc impossible de déterminer les éléments de la séquence responsables de la divergence mécanistique. Par ailleurs, les résultats obtenus chez *S. enterica* en présence de bicyclomycine ne semblent pas expliquer entièrement la régulation observée (Hollands et al., 2012). En effet, dans la même étude les auteurs ont montré que le riborégulateur *ribB* chez *E. coli* est dépendant de la protéine Rho. Les résultats obtenus pour

ribB sont plus clairs que ceux obtenus pour *mgtA* ce qui peut suggérer que le riborégulateur *mgtA* n'est pas entièrement dépendant de la protéine Rho. De plus, en absence de Rho il est possible d'observer *in vitro* la terminaison de la transcription prématurée (Cromie et al., 2006) ce qui confirme que la protéine Rho n'est pas nécessaire bien qu'elle aide à terminer plus efficacement la transcription. Il est donc envisageable que les riborégulateurs *mgtA* chez *E. coli* et *S. enterica* partagent un même mécanisme de régulation de la transcription, avec l'aide de Rho uniquement chez *S. enterica*.

Chez *S. enterica*, bien que les études soient en désaccord sur plusieurs points, elles s'accordent sur un point : le peptide est traduit en fort Mg^{2+} seulement. Les résultats présentés dans cette thèse chez *E. coli* montrent l'inverse. De plus, une étude sommaire du riborégulateur de *S. enterica* suggère que le peptide de *S. enterica* est en fait traduit en faible Mg^{2+} . En réalité, la différence entre les deux études ne provient pas des résultats, mais plutôt des interprétations. Par exemple, dans l'étude de Park et al. (2010), les auteurs interprètent l'augmentation de l'expression de *mgtA* suite à l'ajout de tétracycline par une activation de *mgtA* en absence de traduction du peptide. Or, la tétracycline bloque l'élongation de la traduction bloquant ainsi le ribosome en cours de traduction sur le riborégulateur, imitant un *calage* du ribosome, responsable de l'augmentation de l'expression de *mgtA* selon les résultats obtenus au chapitre 4.

6.5. RIBORÉGULATEUR OU PEPTIDE LEADER?

La présence d'un peptide leader dans un riborégulateur est unique au gène *mgtA*. Gall et al. (2016) proposent que la région 5' du riborégulateur *mgtA* ne contient pas de riborégulateur, mais uniquement un peptide leader et que le Mg^{2+} n'influence l'expression de *mgtA* qu'en contrôlant la vitesse de traduction de *mgtL*. Un riborégulateur est un élément d'ARN capable de lier spécifiquement un métabolite ou un ion, entraînant un changement de conformation de

l'ARN responsable de la régulation de l'expression génique. Les peptides leaders sont des cadres de lecture localisés en amont du ou des gènes régulés et dont la traduction est contrôlée par un signal cellulaire tel que le manque d'un acide aminé ou la présence d'un antibiotique. Le signal cellulaire contrôle alors directement l'activité du ribosome. Déterminer si la région 5' de *mgtA* contient à la fois un riborégulateur et un peptide leader ou uniquement un peptide leader revient à déterminer le rôle exact de l'ion Mg^{2+} dans le mécanisme de régulation. Si effectivement son rôle est uniquement de réguler l'activité du ribosome en ralentissement l'élongation de la traduction, la notion de riborégulateur ne s'applique pas. Mais si le Mg^{2+} induit un changement de conformation de l'ARN qui est responsable de la régulation de la traduction du peptide alors il est possible de dire que l'expression de *mgtA* est régulée par un riborégulateur Mg^{2+} via un peptide leader. Dans le chapitre 3, une étude structurale du riborégulateur *in vitro* en absence de ribosome montre quelques changements de structure en fonction de la concentration en Mg^{2+} . Néanmoins, ces changements ne sont pas extrêmes et la région SD de *mgtL* semble inaccessible quelle que soit la concentration de Mg^{2+} . Par ailleurs, contrairement à d'autres riborégulateurs incluant le riborégulateur Mg^{2+} *mgtE*, le site de liaison de l'ion Mg^{2+} n'a pas été déterminé. Le Mg^{2+} se liant à l'ARN de manière générale il est envisageable qu'il n'y ait pas un site de liaison précis, mais qu'il y ait plusieurs ions Mg^{2+} le long de l'ARN. Néanmoins une pochette de haute affinité au Mg^{2+} a été découverte dans la région 5' de *mgtA* chez *Yersinia enterocolitica* (Korth et Sigel, 2012). Ce constat supporte l'hypothèse que le Mg^{2+} a un rôle direct sur l'ARN et non uniquement sur la vitesse de traduction. Bien qu'une absence de changement de l'accessibilité du SD de *mgtL* dans le cartographique enzymatique du chapitre 3, il est possible qu'en faible Mg^{2+} la structure inaccessible aux RNases utilisées soit plus facilement défaite par le ribosome que celle formée en fort Mg^{2+} . D'autre part, la fusion hybride *thiM-mgtL-mgtA* montre qu'un changement de structure est suffisant pour contrôler l'expression de *mgtL* et ce indépendamment de la concentration en Mg^{2+} . Si la vitesse de traduction du peptide était nécessaire à la régulation, un changement de structure ne devrait pas être suffisant. L'ensemble de ces observations soutient le modèle du riborégulateur qui utilise un peptide leader plutôt que l'absence de riborégulateur.

6.6. RÉGULATION PAR UN PARN

Une étude récente a identifié un pARN capable de s'hybrider à la région 5' de *mgtA* par l'intermédiaire de la protéine chaperonne Hfq (Melamed et al., 2016). Bien que l'interaction entre les deux ARN ait été établie *in vivo*, l'effet de la liaison du pARN sur la régulation de l'expression *mgtA* reste à déterminer. Une autre équipe a montré que l'absence de la protéine Hfq diminue le niveau d'ARNm de *mgtA* d'au moins 1,5X (Guisbert et al., 2007). Conjointement, ces études suggèrent que l'expression de *mgtA* peut être modulée par la présence d'un pARN, qui proviendrait de la région 3' du gène *ykgH*. La liaison de ce pARN, après le cadre de lecture de *mgtL*, nécessiterait la chaperonne Hfq et stabiliserait le messager. Les conditions de production du pARN ne sont pas connues, ni la fonction du gène *ykgH*. La liaison du pARN pourrait faire partie du mécanisme du riborégulateur ou plus probablement permettrait la régulation de l'expression de *mgtA* à un autre signal cellulaire. La localisation de la région d'interaction directement en 3' du cadre de lecture du peptide leader suggère que celui-ci pourrait avoir un rôle dans le mécanisme de régulation par le pARN.

6.7. DÉGRADATION

Bien qu'un mécanisme de régulation de l'élongation de la transcription mis en évidence par l'expérience de CHIP ait été découvert, il n'est pas improbable que le niveau d'ARNm soit également régulé au niveau de sa dégradation. La demi-vie de l'ARN pleine longueur *mgtA* est très faible, moins d'une minute, ce qui suggère une dégradation active de l'ARN. La RNase E et le dégradosome n'étant pas impliqués, plusieurs RNases pourraient être responsables de cette dégradation. La RNase R et la RNase III sont les principales candidates puisqu'elles ont la particularité de cliver les ARN structurés, ce dont ne sont pas capables toutes les RNases. Toutes deux ont par ailleurs comme cofacteur le Mg^{2+} , ce qui pourrait expliquer qu'en fort Mg^{2+} il y ait plus de dégradation qu'en faible. D'autre part, la RNase R est plus fortement

exprimée en phase stationnaire qu'en phase exponentielle ce qui pourrait expliquer la régulation est sensible à la phase de croissance (Andrade et al., 2006).

6.8. RÉGULATION DE LA TRADUCTION

L'étude présentée dans cette thèse s'est concentrée sur l'effet du Mg^{2+} sur le niveau d'ARNm, mais comme tous les riborégulateurs connus de *E. coli*, le riborégulateur *mgtA* est supposé utiliser un mécanisme traductionnel (Barrick et Breaker, 2007). La deuxième structure en 3' séquestre le SD de *mgtA* en fort Mg^{2+} et les données obtenues ne permettent pas de conclure pour la conformation présente en faible Mg^{2+} . Bien que la régulation de la traduction ne soit pas nécessaire étant donné que la région codante de *mgtA* ne sera que très peu transcrite, la régulation observée dans le système minimal de traduction *in vitro* suggère qu'il existe une régulation traductionnelle. La combinaison d'un mécanisme traductionnel et d'un mécanisme transcriptionnel n'est pas unique. En effet, le riborégulateur *lysC* chez *E. coli* par exemple utilise un mécanisme traductionnel, un mécanisme transcriptionnel dépendant de Rho ainsi qu'un mécanisme de dégradation par la RNase E (résultats du laboratoire). Le rôle de la régulation multi-mécanistique reste sujet à discussion. La combinaison de mécanismes agissant à différents niveaux pourrait permettre une régulation plus efficace, le mécanisme traductionnel servant en dernier recours, dans le cas où les autres mécanismes ne performant pas complètement. Les différents mécanismes pourraient également répondre à différentes gammes de concentration de ligands, permettant ainsi une régulation fine de l'expression en fonction des conditions. Enfin, le mécanisme traductionnel ne nécessitant aucune protéine, il pourrait être un vestige d'un « monde ARN ».

6.9. MODÈLE PROPOSÉ

Les résultats obtenus aux chapitres 3 et 4 ont permis de monter un modèle du mécanisme de régulation de la région 5' de *mgtA*, présenté à la Figure 35. Ce modèle explique les rôles respectifs du riborégulateur et du peptide leader, mais ne précise pas par quel mécanisme la présence de ribosome *calé* influence l'élongation de la transcription. La possibilité que le ribosome protège un site *rut* ou la formation d'un terminateur intrinsèque a été écartée dans le chapitre 3. Une étude récente a montré qu'une longue pause de la PdARN peut résulter en un « embouteillage » transcriptionnel bloquant ainsi la transcription (Yuzenkova et al., 2014). Les résultats de ChIP-qPCR obtenus le long du gène *mgtA* montrent qu'il y a une accumulation de la PdARN dans la région 5' de *mgtA* (le pic observé à la Figure 20B). Il est donc possible que la pause de la PdARN au nucléotide C214 observée à la Figure 21A soit responsable d'un tel « embouteillage », résolu par la présence de ribosomes sur *mgtL*, en amont de la pause, permettant de débloquer la PdARN (Figure 57). Cette hypothèse explique également pourquoi la seule région du peptide leader n'est pas suffisante pour réguler l'expression de *mgtA* : la présence de la longue pause transcriptionnelle au C214 est également nécessaire.

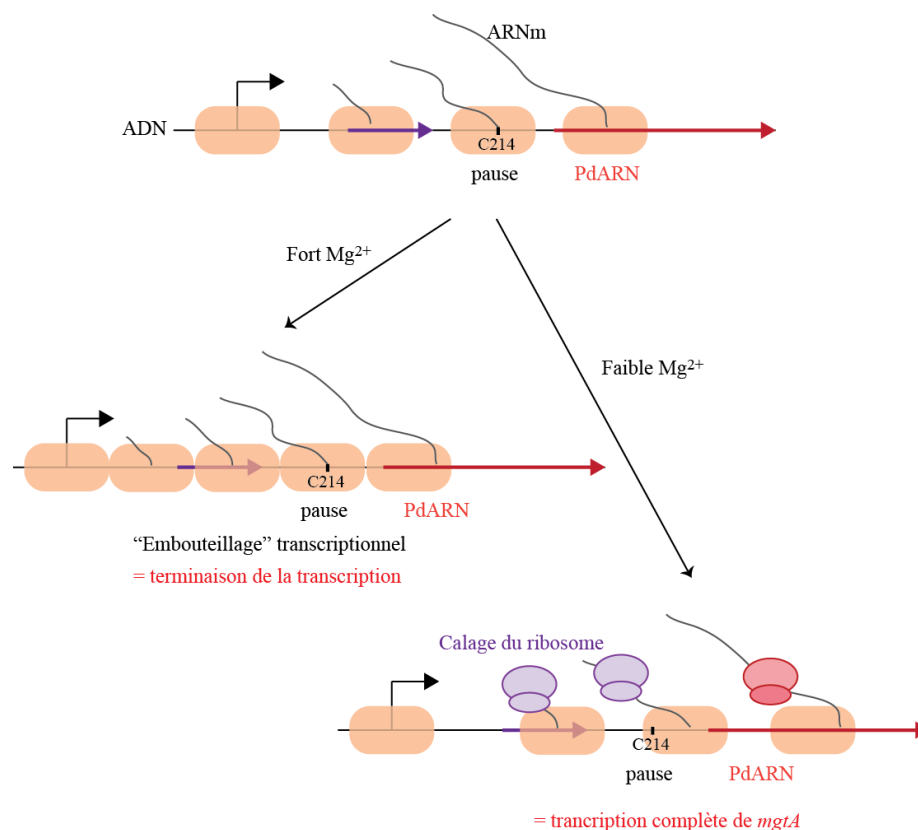


Figure 57. Modèle de l'« embouteillage » transcriptionnel. En absence de ribosome sur le cadre de lecture *mgtL* (violet), un « embouteillage » de PdARN se forme menant à la terminaison prématurée de la transcription. En faible Mg^{2+} , le *calage* du ribosome sur *mgtL* permet à la PdARN de sortir de sa pause évitant ainsi l'« embouteillage ».

Une autre étude a montré que la protéine SuhB interagit avec le ribosome et a observé un *calage* du ribosome sur un peptide leader d'un ARNm de *Pseudomonas aeruginosa* dans un mutant de celle-ci (Shi et al., 2015). En conséquence de ce *calage*, la structure secondaire de l'ARN est affectée, empêchant la formation d'un terminateur de transcription. Ce modèle ressemble à celui des peptides leaders connus et implique un nouvel acteur : la protéine SuhB. Bien qu'il soit peu probable que la protéine SuhB soit impliquée dans le mécanisme de régulation de *mgtA* (résultats du laboratoire du Pr Joseph T. Wade), il est envisageable qu'une autre protéine puisse influencer le calage du ribosome et donc l'expression de *mgtA*.

Ainsi une combinaison de ces deux mécanismes pourrait expliquer le mécanisme de régulation du riborégulateur *mgtA*. Une étude des facteurs protéiques associés à la région 5' de *mgtA* permettrait alors d'identifier les acteurs de ce mécanisme atypique.

6.10. DÉCOUVERTE DE NOUVEAUX GÈNES RÉGULÉS PAR LE Mg^{2+}

La régulation de l'expression génique par le Mg^{2+} est étudiée depuis plusieurs décennies, il est donc surprenant de découvrir de nouveaux candidats au régulon Mg^{2+} . Le *ribosome profiling* est une technique développée récemment qui permet d'observer l'expression génique au niveau génomique. La technique utilisée jusqu'alors était le séquençage ARN haut débit qui se limite à évaluer le niveau d'ARNm et ne prend donc pas en considération le taux de traduction de ces ARNm. Il est donc possible qu'en utilisant le profilage ribosomique plutôt que le séquençage ARN, de nouveaux gènes régulés par le Mg^{2+} aient été découverts, si la régulation se fait principalement au niveau de la traduction. De plus, le profilage ribosomique a été réalisé dans de nouvelles conditions de croissance et de Mg^{2+} . Les études précédentes utilisaient en effet des concentrations élevées de Mg^{2+} (par exemple, dans le milieu riche LB (Luria Broth), avec ou sans ajout de 30 mM de Mg^{2+} , le LB contenant déjà 50 μ M de Mg^{2+} , concentration élevée pour le riborégulateur *mgtA*). Par ailleurs, les méthodes publiées pour le profilage ribosomique suggèrent fortement de se limiter à la phase exponentielle de croissance, or les résultats obtenus au chapitre 5 proviennent de cultures en début de phase stationnaire, ce qui n'a jamais été fait. Sachant que pour l'expression de *mgtA* la phase de croissance est importante, il est possible que certains gènes régulés par le Mg^{2+} n'aient pas été découverts s'ils ne le sont qu'en phase stationnaire. D'autre part, une partie des études s'est concentrée sur la régulation par le Mg^{2+} dépendante du régulateur PhoP, or il est possible que certains gènes répondent au Mg^{2+} de manière indépendante de PhoP, d'autant plus que celui-ci peut répondre à d'autres signaux que le Mg^{2+} , tels que le pH ou les peptides antimicrobiens cationiques.

Parmi les gènes les plus significativement régulés par le Mg^{2+} , deux transporteurs de Mg^{2+} ont été identifiés : CorA et YbeX. Le lien entre la fonction de ces gènes et le signal régulateur suggère que ces deux gènes sont effectivement régulés par le Mg^{2+} et que ce n'est pas un artéfact bien qu'aucune confirmation expérimentale n'ait pu être obtenue dans cette étude. Bien que CorA soit considéré comme le transporteur basal et que son expression est souvent considérée comme constitutive, une étude a montré une régulation par le Mg^{2+} en phase stationnaire (Papp-Wallace et Maguire, 2008). Des discussions avec le Pr Maude Guillier à Paris, ont confirmés l'hypothèse d'une régulation par le Mg^{2+} au niveau post-transcriptionnel, indépendamment du promoteur, laissant suspecter la présence d'un autre riborégulateur Mg^{2+} dans le 5' de *corA*.

Deux gènes ont pu être confirmés comme faisant partie du régulon Mg^{2+} dans cette étude : les gènes *ynfB* et *ompX*. Ces deux gènes n'ont pas de fonction connue à ce jour, ce qui peut expliquer le manque d'information sur leur régulation. La régulation par le Mg^{2+} signifie probablement que leur fonction a un lien avec le Mg^{2+} .

De manière intéressante ces deux gènes codent pour des protéines de la membrane externe et quatre autres gènes (*corA*, *ybeX*, *gsiA* et *mfaF*) codent pour des protéines de la membrane interne, les protéines membranaires comptant ainsi pour 67 % des gènes candidats. La carence en Mg^{2+} semble ainsi résulter en un réarrangeant la composition de la membrane, permettant probablement de protéger la cellule et de maintenir l'homéostasie.

6.11. CONTRIBUTIONS

L'étude présentée aux chapitres 3 et 4 a permis d'apporter de nouveaux éléments pour aider à la compréhension du mécanisme de régulation du riborégulateur *mgtA*. En plus de montrer que

le riborégulateur est fonctionnel chez *E. coli*, l'étude a montré que le peptide leader était exprimé en faible Mg^{2+} et non en fort Mg^{2+} comme suggéré dans les études précédentes. Elle a également mis en évidence l'importance du *calage* du ribosome sur *mgtL* sur la régulation. Ces deux chapitres sont le sujet d'un article en cours de rédaction.

Afin d'étudier le rôle de la traduction du peptide sur l'expression de *mgtA*, un outil a été développé. Cet outil consiste à utiliser un riborégulateur traductionnel pour réguler la traduction d'un peptide leader et, lorsqu'utilisé en fusion avec un gène rapporteur, il permet d'évaluer l'impact de la traduction d'un peptide leader d'intérêt sur l'expression du gène en aval. En plus d'apporter des nouvelles informations sur le peptide leader *mgtL*, cet outil a également servi à montrer que la traduction du peptide *corL* est importante pour l'expression de *corA*. Ce dernier résultat suggère que l'outil développé pour un peptide leader peut être adapté à d'autres peptides leaders. L'avantage de cet outil est de garder le contexte génétique du peptide (pas de mutation, pas de troncature). L'inconvénient est qu'il est limité à la régulation de l'initiation de la traduction du peptide.

L'expérience de profilage ribosomique a généré de nombreuses données. Une première analyse est présentée dans le chapitre 5, mais il est possible d'extraire encore plus d'informations. Il serait par exemple intéressant de comparer ces données avec celles de la littérature obtenues en de phase exponentielle pour identifier des gènes dont l'expression est dépendante de la phase de croissance. Le but premier de cette expérience était d'étudier l'expression du peptide dans son contexte endogène. En effet, la technique utilisée jusqu'à présent, une fusion traductionnelle avec un gène rapporteur, comporte le gros désavantage de perdre la structure de l'ARN, qui peut jouer un rôle majeur dans l'expression de celui-ci. La technique a ainsi permis de mettre en évidence une régulation de l'expression du peptide inverse à celle supposée jusqu'à présent. Le profilage ribosomique pour ces raisons est une technique intéressante pour l'étude des peptides leaders en général. Cette technique a par

ailleurs permis de découvrir de nouveaux gènes régulés par le Mg^{2+} , ayant jusqu'à présent passé inaperçu en séquençage ARN. Cette technique a également permis de redéfinir le cadre de lecture du gène *ybjX* et pourra probablement aider à la définition d'autres de cadres de lecture.

BIBLIOGRAPHIE

Aguilera, A., et García-Muse, T. (2012). R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol. Cell* *46*, 115-124.

Alifano, P., Fani, R., Liò, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M.S., et Bruni, C.B. (1996). Histidine biosynthetic pathway and genes: structure, regulation, and evolution. *Microbiol. Rev.* *60*, 44-69.

Altman, S., Wesolowski, D., Guerrier-Takada, C., et Li, Y. (2005). RNase P cleaves transient structures in some riboswitches. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 11284-11289.

Ames, T.D., et Breaker, R.R. (2011). Bacterial aptamers that selectively bind glutamine. *RNA Biol.* *8*, 82-89.

Ames, T.D., Rodionov, D.A., Weinberg, Z., et Breaker, R.R. (2010). A Eubacterial Riboswitch Class that Senses the Coenzyme Tetrahydrofolate. *Chem. Biol.* *17*, 681–685.

Andrade, J.M., Cairrão, F., et Arraiano, C.M. (2006). RNase R affects gene expression in stationary phase: Regulation of ompA. *Mol. Microbiol.* *60*, 219-228.

Anupama, K., Leela, J.K., et Gowrishankar, J. (2011). Two pathways for RNase E action in *Escherichia coli in vivo* and bypass of its essentiality in mutants defective for Rho-dependent

transcription termination. *Mol. Microbiol.* 82, 1330-1348.

Apirion, D., et Lassar, A.B. (1978). Lethal Mutant of *Escherichia* Affects the Processing of Ribosomal RNA. *J. Biol. Chem.* 253, 1738-1742.

Arraiano, C.M., Andrade, J.M., Domingues, S., Guinote, I.B., Malecki, M., Matos, R.G., Moreira, R.N., Pobre, V., Reis, F.P., Saramago, M., et al. (2010). The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.* 34, 883-923.

Artsimovitch, I., et Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci.* 97, 7090-7095.

Awano, N., Rajagopal, V., Arbing, M., Patel, S., Hunt, J., Inouye, M., et Phadtare, S. (2010). *Escherichia coli* RNase R has dual activities, helicase and RNase. *J. Bacteriol.* 192, 1344-1352.

Bader, M.W., Navarre, W.W., Shiau, W., Nikaido, H., Frye, J.G., McClelland, M., Fang, F.C., et Miller, S.I. (2003). Regulation of *Salmonella typhimurium* virulence gene expression by cationic antimicrobial peptides. *Mol. Microbiol.* 50, 219-230.

Baker, C.S., Morozov, I., Suzuki, K., Romeo, T., et Babitzke, P. (2002). CsrA regulates glycogen biosynthesis by preventing translation of glgC in *Escherichia coli*. *Mol. Microbiol.* 44, 1599-1610.

Baker, J.L., Sudarsan, N., Weinberg, Z., Roth, A., Stockbridge, R.B., et Breaker, R.R. (2012). Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science*. 335, 233-235.

Balaban, N., et Novick, R.P. (1995). Translation of RNAlII, the *Staphylococcus aureus* agr regulatory RNA molecule, can be activated by a 3'-end deletion. *FEMS Microbiol. Lett.* 133, 155-161.

Banerjee, S., Chalissery, J., Bandey, I., et Sen, R. (2006). Rho-dependent transcription termination: more questions than answers. *J. Microbiol.* 44, 11-22.

Barchiesi, J., Castelli, M.E., Soncini, F.C., et Vescovi, E.G. (2008). *mgtA* expression is induced by Rob overexpression and mediates a *Salmonella enterica* resistance phenotype. *J. Bacteriol.* 190, 4951-4958.

Barrick, J.E., et Breaker, R.R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* 8, R239.

Barrick, J.E., Corbino, K.A., Winkler, W.C., Nahvi, A., Mandal, M., Collins, J., Lee, M., Roth, A., Sudarsan, N., Jona, I., et al. (2004). New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci.* 101, 6421-6426.

Bastet, L., Dubé, A., Massé, E., et Lafontaine, D.A. (2011). New insights into riboswitch regulation mechanisms. *Mol. Microbiol.* 80, 1148-1154.

Bastet, L., Chauvier, A., Singh, N., Lussier, A., Wade, J.T., Lamontagne, A., Pr, K., Mass, E., et Lafontaine, A. (2017). Translational control and Rho-dependent transcription termination are intimately linked in riboswitch. *45*, 7474-7486.

Belasco, J.G. (2010). All things must pass: Contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat. Rev. Mol. Cell Biol.* *11*, 467-478.

Benjamini, Y., et Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* *57*, 289-300.

Bernstein, J. a, Lin, P.-H., Cohen, S.N., et Lin-Chao, S. (2004). Global analysis of Escherichia coli RNA degradosome function using DNA microarrays. *Proc. Natl. Acad. Sci.* *101*, 2758-2763.

Bischoff, L., Berninghausen, O., et Beckmann, R. (2014). Molecular basis for the ribosome functioning as an L-tryptophan sensor. *Cell Rep.* *9*, 469-475.

Bolger, A.M., Lohse, M., et Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114-2120.

Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., et Rösch, P. (2010). A NusE: NusG complex links transcription and translation. *Science.* *328*, 501-504.

Butkus, M.E., Prundeanu, L.B., et Oliver, D.B. (2003). Translocon « Pulling » of Nascent

SecM Controls the Duration of Its Translational Pause and Secretion-Responsive secA Regulation. *J. Bacteriol.* *185*, 6719-6722.

Calin-Jageman, I., et Nicholson, A.W. (2003). RNA structure-dependent uncoupling of substrate recognition and cleavage by *Escherichia coli* ribonuclease III. *Nucleic Acids Res.* *31*, 2381-2392.

Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., et Darst, S.A. (2001). Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* *104*, 901-912.

Caron, M., Bastet, L., Lussier, A., Simoneau-roy, M., Massé, E., et Lafontaine, D.A. (2012). Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc. Natl. Acad. Sci.* *109*.

Caron, M.P., Lafontaine, D.A., et Massé, E. (2010). Small RNA-mediated regulation at the level of transcript stability. *RNA Biol.* *7*, 140-144.

Carpousis, A.J., Luisi, B.F., et McDowall, K.J. (2009). Endonucleolytic Initiation of mRNA Decay in *Escherichia coli*. *Prog. Mol. Biol. Transl. Sci.* *85*, 91-135.

Casadesus, J., et Low, D. (2006). Epigenetic Gene Regulation in the Bacterial World. *Microbiol. Mol. Biol. Rev.* *70*, 830-856.

Cavanagh, A.T., et Wassarman, K.M. (2014). 6S RNA, a Global Regulator of Transcription in *Escherichia coli* , *Bacillus subtilis* , and Beyond. *Annu. Rev. Microbiol.* 68, 45-60.

Chadani, Y., Niwa, T., Izumi, T., Sugata, N., Nagao, A., Suzuki, T., Chiba, S., Ito, K., et Taguchi, H. (2017). Intrinsic Ribosome Destabilization Underlies Translation and Provides an Organism with a Strategy of Environmental Sensing. *Mol. Cell* 68, 528-539.e5.

Chang, D.-E., Smalley, D.J., et Conway, T. (2002). Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. *Mol. Microbiol.* 45, 289-306.

Chauvier, A., Picard-Jean, F., Berger-Dancause, J.-C., Bastet, L., Naghdi, M.R., Dubé, A., Turcotte, P., Perreault, J., et Lafontaine, D.A. (2017). Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat. Commun.* 8, 13892.

Cheng, Z.F., et Deutscher, M.P. (2005). An important role for RNase R in mRNA decay. *Mol. Cell* 17, 313-318.

Chiba, S., Lamsa, A., et Pogliano, K. (2009). A ribosome-nascent chain sensor of membrane protein biogenesis in *Bacillus subtilis*. *EMBO J.* 28, 3461-3475.

Chopra, I., et Roberts, M. (2001). Tetracycline Antibiotics : Mode of Action , Applications , Molecular Biology , and Epidemiology of Bacterial Resistance. *Microbiol. Mol. Biol. Rev.* 65, 232-260.

Ciampi, M.S. (2006). Rho-dependent terminators and transcription termination. *Microbiology* *152*, 2515-2528.

Corbin, R.W., Paliy, O., Yang, F., Shabanowitz, J., Platt, M., Lyons, C.E., Root, K., McAuliffe, J., Jordan, M.I., Kustu, S., et al. (2003). Toward a protein profile of *Escherichia coli*: Comparison to its transcription profile. *Proc. Natl. Acad. Sci.* *100*, 9232-9237.

Cromie, M.J., et Groisman, E.A. (2010). Promoter and riboswitch control of the Mg²⁺ transporter MgtA from *Salmonella enterica*. *J. Bacteriol.* *192*, 604-607.

Cromie, M.J., Shi, Y., Latifi, T., et Groisman, E.A. (2006). An RNA Sensor for Intracellular Mg²⁺. *Cell* *125*, 71-84.

Culver, G.M., et Kirthi, N. (2008). Assembly of the 30S Ribosomal Subunit. *EcoSal Plus* *3*.

D'Heygère, F., Rabhi, M., et Boudvillain, M. (2013). Phyletic distribution and conservation of the bacterial transcription termination factor rho. *Microbiol. (United Kingdom)* *159*, 1423-1436.

Dambach, M., Sandoval, M., Updegrove, T.B., Anantharaman, V., Aravind, L., Waters, L.S., et Storz, G. (2015). The ubiquitous yybP-ykoY riboswitch is a manganese- responsive regulatory element. *Mol. Cell* *57*, 1099-1109.

Dann III, C.E., Wakeman, C.A., Sieling, C.L., Baker, S.C., Irnov, I., et Winkler, W.C. (2007).

Structure and Mechanism of a Metal-Sensing Regulatory RNA. *Cell* 130, 878-892.

Dar, D., et Sorek, R. (2018). High-resolution RNA 3'-ends mapping of bacterial Rho-dependent transcripts. *Nucleic Acids Res.* 1-9.

Darfeuille, F., Unoson, C., Vogel, J., et Wagner, E.G.H. (2007). An Antisense RNA Inhibits Translation by Competing with Standby Ribosomes. *Mol. Cell* 26, 381-392.

Datsenko, K.A., et Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640-6645.

David, G., Blondeau, K., Schiltz, M., Penel, S., et Lewit-Bentley, A. (2003). YodA from *Escherichia coli* is a Metal-binding, Lipocalin-like Protein. *J. Biol. Chem.* 278, 43728-43735.

Deana, A., et Belasco, J.G. (2005). Lost in translation: The influence of ribosomes on bacterial mRNA decay. *Genes Dev.* 19, 2526-2533.

Deana, A., Celesnik, H., et Belasco, J.G. (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* 451, 355-358.

DeRose, V.J. (2007). Sensing cellular magnesium with RNA. *Nat. Chem. Biol.* 3, 693-694.

Detweiler, C.S., Monack, D.M., Brodsky, I.E., Mathew, H., et Falkow, S. (2003). *virK*, *somA*

and *rscC* are important for systemic *Salmonella enterica* serovar Typhimurium infection and cationic peptide resistance. *Mol. Microbiol.* *48*, 385-400.

Deutscher, M.P. (2006). Degradation of RNA in bacteria: Comparison of mRNA and stable RNA. *Nucleic Acids Res.* *34*, 659-666.

Dorman, C.J. (2013). Co-operative roles for DNA supercoiling and nucleoid-associated proteins in the regulation of bacterial transcription. *Biochem. Soc. Trans.* *41*, 542-547.

Drolet, M., Phoenix, P., Menzel, R., Massé, E., Liu, L.F., et Crouch, R.J. (1995). Overexpression of RNase H partially complements the growth defect of an *Escherichia coli* delta *topA* mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. *Proc. Natl. Acad. Sci. U. S. A.* *92*, 3526-3530.

Dubnau, D., et Nomura, M. (1984). Translational attenuation: The regulation of bacterial resistance to the macrolide-lincosamide-streptogramin b antibiotic. *Crit. Rev. Biochem. Mol. Biol.* *16*, 103-132.

Dutta, T., et Srivastava, S. (2018). Small RNA-mediated regulation in bacteria: A growing palette of diverse mechanisms. *Gene* *656*, 60-72.

Duval, M., Korepanov, A., Fuchsbaauer, O., Fechter, P., Haller, A., Fabbretti, A., Choulier, L., Micura, R., Klaholz, B.P., Romby, P., et al. (2013). *Escherichia coli* Ribosomal Protein S1 Unfolds Structured mRNAs Onto the Ribosome for Active Translation Initiation. *PLoS Biol.*

11, 12-14.

Duvall, E.J., Williams, D.M., Lovett, P.S., Rudolph, C., Vasantha, N., et Guyer, M. (1983). Chloramphenicol-inducible gene expression in *Bacillus subtilis*. *Gene* 24, 171-177.

Ekiert, D.C., Bhabha, G., Isom, G.L., Greenan, G., Ovchinnikov, S., Henderson, I.R., Cox, J.S., et Vale, R.D. (2017). Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. *Cell* 169, 273-285.e17.

Ferré-D'Amaré, A.R., et Scott, W.G. (2010). Small self-cleaving ribozymes. *Cold Spring Harb. Perspect. Biol.* 2, 1-10.

Franch, T., Gultyaev, A.P., et Gerdes, K. (1997). Programmed cell death by hok/sok of plasmid R1: Processing at the hok mRNA 3'-end triggers structural rearrangements that allow translation and antisense RNA binding. *J. Mol. Biol.* 273, 38-51.

Freistroffer, D. V., Pavlov, M.Y., MacDougall, J., Buckingham, R.H., et Ehrenberg, M. (1997). Release factor RF3 in *E. coli* accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. *EMBO J.* 16, 4126-4133.

Froschauer, E.M., Kolisek, M., Dieterich, F., Schweigel, M., et Schweyen, R.J. (2004). Fluorescence measurements of free [Mg²⁺] by use of mag-fura 2 in *Salmonella enterica*. *FEMS Microbiol. Lett.* 237, 49-55.

Gall, A.R., Datsenko, K.A., Figueroa-Bossi, N., Bossi, L., Masuda, I., Hou, Y.-M., et Csonka, L.N. (2016). Mg^{2+} regulates transcription of *mgtA* in *Salmonella* Typhimurium via translation of proline codons during synthesis of the MgtL peptide. *Proc. Natl. Acad. Sci.* *113*, 201612268.

Galluppi, G.R., et Richardson, J.P. (1980). ATP-induced changes in the binding of RNA synthesis termination protein rho to RNA. *J. Mol. Biol.* *138*, 513-539.

Garcia Vescovi, E., Ayala, Y.M., Di Cera, E., et Groisman, E.A. (1997). Characterization of the Bacterial Sensor Protein PhoQ. *J. Biol. Chem.* *272*, 1440-1443.

Gardner, J. (1979). Regulation of the threonine operon: tandem threonine and isoleucine codons in the control region and translational control of transcription termination. *Proc. Natl. Acad. ...* *76*, 1706-1710.

Gemmill, R.M., Wessler, S.R., Keller, E.B., et Calvo, J.M. (1979). *leu* operon of *Salmonella typhimurium* is controlled by an attenuation mechanism. *Proc. Natl. Acad. Sci.* *76*, 4941-4945.

Gesteland, R.F. (1966). Unfolding of *Escherichia coli* ribosomes by removal of magnesium. *J. Mol. Biol.* *18*, IN12-IN14.

Ghosh, S., et Deutscher, M.P. (1999). Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 4372-4377.

Gibson, M.M., Bagga, D.A., Miller, C.G., et Maguire, M.E. (1991). Magnesium transport in *Salmonella typhimurium*: the influence of new mutations conferring Co²⁺ resistance on the CorA Mg²⁺ transport system. *Mol. Microbiol.* 5, 2753-2762.

Gold, L. (1988). Posttranscriptional regulatory mechanisms in *ESCHERICHIA COLI*. *Annu. Rev. Biochem.* 57, 199-233.

Gollnick, P. (1994). Regulation of the *Bacillus subtilis* trp operon by an RNA-binding protein. *Mol. Microbiol.* 11, 991-997.

Gollnick, P., et Babitzke, P. (2002). Transcription attenuation. *Biochim. Biophys. Acta* 1577, 240-250.

Gowrishankar, J., et Harinarayanan, R. (2004). Why is transcription coupled to translation in bacteria? *Mol. Microbiol.* 54, 598-603.

Graf, M., Mardirossian, M., Nguyen, F., Seefeldt, A.C., Guichard, G., Scocchi, M., Innis, C.A., et Wilson, D.N. (2017). Proline-rich antimicrobial peptides targeting protein synthesis. *Nat. Prod. Rep.* 34, 702-711.

Graham, A.I., Hunt, S., Stokes, S.L., Bramall, N., Bunch, J., Cox, A.G., McLeod, C.W., et Poole, R.K. (2009). Severe zinc depletion of *Escherichia coli*: Roles for high affinity zinc binding by ZinT, zinc transport and zinc-independent proteins. *J. Biol. Chem.* 284, 18377-18389.

Grill, S., Gualerzi, C.O., Londei, P., et Bläsi, U. (2000). Selective stimulation of translation of leaderless mRNA by initiation factor 2: Evolutionary implications for translation. *EMBO J.* 19, 4101-4110.

Groisman, E. a (1998). The ins and outs of virulence gene expression : Mg²⁺ as a regulatory signal. *Bioessays* 20, 96-101.

Gruber, T.M., et Gross, C.A. (2003). Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. *Annu. Rev. Microbiol.* 57, 441-466.

Gualerzi, C.O., et Pon, C.L. (2015). Initiation of mRNA translation in bacteria: Structural and dynamic aspects. *Cell. Mol. Life Sci.* 72, 4341-4367.

Guimaraes, J.C., Rocha, M., et Arkin, A.P. (2014). Transcript level and sequence determinants of protein abundance and noise in *Escherichia coli*. *Nucleic Acids Res.* 42, 4791-4799.

Guisbert, E., Rhodius, V.A., Ahuja, N., Witkin, E., et Gross, C.A. (2007). Hfq modulates the σ^E -mediated envelope stress response and the σ^{32} -mediated cytoplasmic stress response in *Escherichia coli*. *J. Bacteriol.* 189, 1963-1973.

Gupta, P., Kannan, K., Mankin, A.S., et Vazquez-Laslop, N. (2013). Regulation of gene expression by macrolide-induced ribosomal frameshifting. *Mol. Cell* 52, 629-642.

Gusarov, I., et Nudler, E. (1999). The mechanism of intrinsic transcription termination. *Mol.*

Cell 3, 495-504.

Guskov, A., et Eshaghi, S. (2012). The Mechanisms of Mg^{2+} and Co^{2+} Transport by the CorA Family of Divalent Cation Transporters (Elsevier).

Hancock, R.E.W. (1984). Alterations in structure of the cell envelope. *Ann. Rev. Microbiol.* 38, 237-264.

Harari, O., Park, S.Y., Huang, H., Groisman, E.A., et Zwir, I. (2010). Defining the plasticity of transcription factor binding sites by deconstructing DNA consensus sequences: The *phoP*-binding sites among gamma/enterobacteria. *PLoS Comput. Biol.* 6, 10.

Harinarayanan, R., et Gowrishankar, J. (2003). Host factor titration by chromosomal R-loops as a mechanism for runaway plasmid replication in transcription termination-defective mutants of *Escherichia coli*. *J. Mol. Biol.* 332, 31-46.

Henkin, T.M., et Grundy, F.J. (2006). Sensing metabolic signals with nascent RNA transcripts: The T box and S box riboswitches as paradigms. *Cold Spring Harb. Symp. Quant. Biol.* 71, 231-237.

Heppell, B., Blouin, S., Dussault, A.M., Mulhbach, J., Ennifar, E., Penedo, J.C., et Lafontaine, D.A. (2011). Molecular insights into the ligand-controlled organization of the SAM-I riboswitch. *Nat. Chem. Biol.* 7, 384-392.

Hollands, K., Proshkin, S., Sklyarova, S., Epshtein, V., Mironov, A., Nudler, E., et Groisman, E. a (2012). Riboswitch control of Rho-dependent transcription termination. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 5376-5381.

Hollands, K., Sevostiyanova, A., et Groisman, E.A. (2014). Unusually long-lived pause required for regulation of a Rho-dependent transcription terminator. *Proc. Natl. Acad. Sci.* *111*, E1999-2007.

Hopkins, F.G., et Cole, S.W. (1903). A CONTRIBUTION TO THE CHEMISTRY OF PROTEIDS. Part II. The Constitution of Tryptophane, and the Action of Bacteria upon it. *J. Physiol.* *29*, 451-466.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., et Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science.* *324*, 218-223.

Ingolia, N.T., Brar, G.A., Rouskin, S., Mcgeachy, A.M., et Weissman, J.S. (2013). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.* *7*, 1534-1550.

Jackson, E.N., et Yaetofsky, C. (1973). The region between the operator and first structural gene of the tryptophan operon of *Escherichia coli* may have a regulatory function. *J. Mol. Biol.* *76*, 89-101.

Jagodnik, J., Chiaruttini, C., et Guillier, M. (2017). Stem-Loop Structures within mRNA Coding Sequences Activate Translation Initiation and Mediate Control by Small Regulatory RNAs. *Mol. Cell* 68, 158-170.e3.

Janosi, L., Ricker, R., et Kaji, A. (1996). Dual functions of ribosome recycling factor in protein biosynthesis: Disassembling the termination complex and preventing translational errors. *Biochimie* 78, 959-969.

Jia, X., Zhang, J., Sun, W., He, W., Jiang, H., Chen, D., et Murchie, A.I.H. (2013). Riboswitch control of aminoglycoside antibiotic resistance. *Cell* 152, 68-81.

Kasai, T. (1974). Regulation of the expression of the histidine operon in *Salmonella typhimurium*. *Nature* 249, 523-527.

Kido, M., Yamanaka, K., Mitani, T., Niki, H., Ogura, T., et Hiraga, S. (1996). RNase E polypeptides lacking a carboxyl-terminal half suppress a mukB mutation in *Escherichia coli*. *J. Bacteriol.* 178, 3917-3925.

Kim, J.N., Roth, A., et Breaker, R.R. (2007). Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2'-deoxyguanosine. *Proc. Natl. Acad. Sci.* 104, 16092-16097.

Kime, L., Jourdan, S.S., Stead, J.A., Hidalgo-Sastre, A., et McDowall, K.J. (2009). Rapid cleavage of RNA by RNase E in the absence of 5' monophosphate stimulation. *Mol. Microbiol.* 76, 590-604.

Knauer, S.H., Rösch, P., et Artsimovitch, I. (2012). Transformation: The next level of regulation. *RNA Biol.* 9, 1418-1423.

Kohler, R., Mooney, R.A., Mills, D.J., Landick, R., et Cramer, P. (2017). Architecture of a transcribing-translating expressome. *Science.* 356, 194-197.

Kohn, H., et Widger, W. (2005). The Molecular Basis for the Mode of Action of Bicyclomycin. *Current Drug Targets – Infectious Disorders.* 5, 273-295.

Komissarova, N., et Kashlev, M. (1997). Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc. Natl. Acad. Sci. U. S. A.* 94, 1755-1760.

Konan, K.V., et Yanofsky, C. (1997). Regulation of the *Escherichia coli tna* Operon : Nascent Leader Peptide Control at the tnaC Stop Codon. *J. Bacteriol.* 179, 1774-1779.

Korth, M.M.T., et Sigel, R.K.O. (2012). Unusually high-affinity Mg²⁺ binding at the AU-rich sequence within the antiterminator hairpin of a Mg²⁺ riboswitch. *Chem. Biodivers.* 9, 2035-2049.

Kortmann, J., et Narberhaus, F. (2012). Bacterial RNA thermometers: Molecular zippers and switches. *Nat. Rev. Microbiol.* 10, 255-265.

Kriner, M.A., et Groisman, E.A. (2015). The Bacterial Transcription Termination Factor Rho

Coordinates Mg^{2+} Homeostasis with Translational Signals. *J. Mol. Biol.* 427, 3834-3849.

Lalaouna, D., Prévost, K., Eyraud, A., et Massé, E. (2017). Identification of unknown RNA partners using MAPS. *Methods* 117, 28-34.

Landick, R., Wang, D., et Chan, C.L. (1996). Quantitative analysis of transcriptional pausing by *Escherichia coli* RNA polymerase: his leader pause site as paradigm. *Methods Enzymol.* 274, 334-353.

Langmead, B., Trapnell, C., Pop, M., et Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10.

Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Nayak, D., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R., et Weissman, J.S. (2014). A pause sequence enriched at translation start sites drives transcription dynamics *in vivo*. *Science.* 344, 1042-1047.

Lawther, R.P., et Hatfield, G.W. (1980). Multivalent translational control of transcription termination at attenuator of *ilvGEDA* operon of *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. U. S. A.* 77, 1862-1866.

Lee, D.J., Minchin, S.D., et Busby, S.J.W. (2012). Activating Transcription in Bacteria. *Annu. Rev. Microbiol.* 66, 125-152.

Lemay, J.F., Desnoyers, G., Blouin, S., Heppell, B., Bastet, L., St-Pierre, P., Massé, E., et

Lafontaine, D.A. (2011). Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms. *PLoS Genet.* 7.

Li, S., et Breaker, R.R. (2013). Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing. *Nucleic Acids Res.* 41, 3022-3031.

Li, G.W., Oh, E., et Weissman, J.S. (2012). The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484, 538-541.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., et Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.

Lim, B., Sim, S., Sim, M., Kim, K., Jeon, O., Lee, Y., Ha, N., et Lee, K. (2012). RNase III Controls the Degradation of corA mRNA in *Escherichia coli*. *J. Bacteriol.* 194, 2214-2220.

Lin, Lin-Chao, S., Lin, P.-H., et Lin-Chao, S. (2005). RhlB helicase rather than enolase is the beta-subunit of the *Escherichia coli* polynucleotide phosphorylase (PNPase)-exoribonucleolytic complex. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16590-16595.

Lopez-Campistrous, A., Semchuk, P., Burke, L., Palmer-Stone, T., Brokx, S.J., Broderick, G., Bottorff, D., Bolch, S., Weiner, J.H., et Ellison, M.J. (2005). Localization, Annotation, and Comparison of the *Escherichia coli* K-12 Proteome under Two States of Growth. *Mol. Cell.*

Proteomics 4, 1205-1209.

Love, M.I., Huber, W., et Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1-21.

Lovett, P.S. (1996). Translation attenuation regulation of chloramphenicol resistance in bacteria - A review. *Gene* 179, 157-162.

Lu, J., et Deutsch, C. (2008). Electrostatics in the Ribosomal Tunnel Modulate Chain Elongation Rates. *J. Mol. Biol.* 384, 73-86.

Lu, J., Kobertz, W.R., et Deutsch, C. (2007a). Mapping the Electrostatic Potential within the Ribosomal Exit Tunnel. *J. Mol. Biol.* 371, 1378-1391.

Lu, P., Vogel, C., Wang, R., Yao, X., et Marcotte, E.M. (2007b). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* 25, 117-124.

Lusk, J.E., Williams, R.J.P., et Kennedy, E.P. (1968). Magnesium and the growth of *Escherichia coli*. *J. Biol. Chem.* 243, 2618-2624.

Maguire, M.E. (2006). The structure of CorA: a Mg^{2+} -selective channel. *Curr. Opin. Struct. Biol.* 16, 432-438.

Maier, T., Schmidt, A., Güell, M., Kühner, S., Gavin, A.C., Aebersold, R., et Serrano, L. (2011). Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.* 7, 1-12.

Mandal, M., et Breaker, R.R. (2004). Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. Mol. Biol.* 11, 29-35.

Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., et Breaker, R.R. (2003). Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* 113, 577-586.

Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L., et Breaker, R.R. (2004). A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science*. 306, 1477.

Martínez-Antonio, A., Janga, S.C., et Thieffry, D. (2008). Functional organisation of *Escherichia coli* transcriptional regulatory network. *J. Mol. Biol.* 381, 238-247.

Massé, E., et Gottesman, S. (2002). A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 4620-4625.

Massé, E., Escorcía, F.E., et Gottesman, S. (2003). Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli.* 2374-2383.

McCarthy, B.J. (1962). The effects of magnesium starvation on the ribosome content of *Escherichia coli*. *Biochim. Biophys. Acta* 55, 880-889.

McClelland, M. (2000). Comparison of the *Escherichia coli* K-12 genome with sampled genomes of a *Klebsiella pneumoniae* and three *Salmonella enterica* serovars, Typhimurium, Typhi and Paratyphi. *Nucleic Acids Res.* 28, 4974-4986.

McDowall, K.J., Hernandez, R.G., Lin-Chao, S., et Cohen, S.N. (1993). The *ams*-1 and *rne*-3071 temperature-sensitive mutations in the *ams* gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the *Escherichia coli mre* locus. *J. Bacteriol.* 175, 4245-4249.

McDowall, K.J., Lin-Chao, S., et Cohen, S.N. (1994). A + U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* 269, 10790-10796.

Mecsas, J., Rouviere, P.E., Erickson, J.W., Donohue, T.J., et Gross, C.A. (1993). The activity of sigma E, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes Dev* 7, 2618-2628.

Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y.E., Reiss, N., Bar, A., Altuvia, Y., Argaman, L., et Margalit, H. (2016). Global Mapping of Small RNA-Target Interactions in Bacteria. *Mol. Cell* 63, 884-897.

Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juárez, K., Contreras-Moreira, B., et al. (2009). Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. PLoS One 4.

Meng, W., et Nicholson, A.W. (2008). Heterodimer-based analysis of subunit and domain contributions to double-stranded RNA processing by *Escherichia coli* RNase III in vitro. Biochem. J. 410, 39-48.

Michaux, C., Verneuil, N., Hartke, A., et Giard, J.-C. (2014). Physiological roles of small RNA molecules. Microbiology 160, 1007-1019.

Minagawa, S., Ogasawara, H., Kato, A., Yamamoto, K., Eguchi, Y., Oshima, T., Mori, H., Ishihama, A., et Utsumi, R. (2003). Identification and Molecular Characterization of the Mg²⁺ Stimulon of *Escherichia coli*. 185, 3696-3702.

Misra, V.K., et Draper, D.E. (2002). The linkage between magnesium binding and RNA folding 1 Edited by B. Honig. J. Mol. Biol. 317, 507-521.

Mitra, A., Angamuthu, K., Jayashree, H.V., et Nagaraja, V. (2009). Occurrence, divergence and evolution of intrinsic terminators across Eubacteria. Genomics 94, 110-116.

Mohanty, B.K., et Kushner, S.R. (1999). Analysis of the function of *Escherichia coli* poly(A) polymerase I in RNA metabolism. Mol Microbiol 34, 1094-108.

Moine, H., Romby, P., Springer, M., Grunberg-Manago, M., Ebel, J.P., Ehresmann, B., et Ehresmann, C. (1990). *Escherichia coli* threonyl-tRNA synthetase and tRNA^{Thr} modulate the binding of the ribosome to the translational initiation site of the ThrS mRNA. *J. Mol. Biol.* 216, 299-310.

Møller, J. V, Juul, B., et le Maire, M. (1996). Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochim. Biophys. Acta* 1286, 1-51.

Moncany, M.L.J., et Kellenberger, E. (1981). High magnesium content of *Escherichia coli* B. *Experientia* 37, 846-847.

Morita, T., Maki, K., et Aiba, H. (2005). RNase E-based ribonucleoprotein complexes: Mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* 19, 2176-2186.

Mumtaz, M.A.S., et Couso, J.P. (2015). Ribosomal profiling adds new coding sequences to the proteome. *Biochem. Soc. Trans.* 43, 1271-1276.

Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., et Breaker, R.R. (2002). Genetic control by a metabolite binding mRNA. *Chem. Biol.* 9, 1043-1049.

Nechooshtan, G., Elgrably-Weiss, M., Sheaffer, A., Westhof, E., et Altuvia, S. (2009). A pH-responsive riboregulator. *Genes Dev.* 23, 2650-2662.

Nelson, J.W., Sudarsan, N., Furukawa, K., Weinberg, Z., Wang, J.X., et Breaker, R.R. (2013). Riboswitches in eubacteria sense the second messenger c-di- AMP. *Nat. Chem. Biol.* 9, 834–839.

Nie, L., Wu, G., et Zhang, W. (2006). Correlation of mRNA expression and protein abundance affected by multiple sequence features related to translational efficiency in *Desulfovibrio vulgaris*: A quantitative analysis. *Genetics* 174, 2229-2243.

Niyogi, S.K., et Datta, A.K. (1975). A Novel Oligoribonuclease of *Escherichia coli*. *The Journal of Biological Chemistry.* 250, 7307-7312.

Nogueira, T., et Springer, M. (2000). Post-transcriptional control by global regulators of gene expression in bacteria. *Curr. Opin. Microbiol.* 3, 154-158.

Nudler, E., et Mironov, A.S. (2004). The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.* 29, 11-17.

Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G., et al. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* 147, 1295-1308.

Papp-Wallace, K.M., et Maguire, M.E. (2008). Regulation of CorA Mg²⁺ channel function affects the virulence of *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* 190, 6509-6516.

Paradis-Bleau, C., Kritikos, G., Orlova, K., Typas, A., et Bernhardt, T.G. (2014). A Genome-Wide Screen for Bacterial Envelope Biogenesis Mutants Identifies a Novel Factor Involved in Cell Wall Precursor Metabolism. *PLoS Genet.* 10.

Park, S.Y., Cromie, M.J., Lee, E.J., et Groisman, E.A. (2010). A Bacterial mRNA Leader that Employs Different Mechanisms to Sense Disparate Intracellular Signals. *Cell* 142, 737-748.

Payandeh, J., Pfoh, R., et Pai, E.F. (2013). The structure and regulation of magnesium selective ion channels. *Biochim. Biophys. Acta - Biomembr.* 1828, 2778-2792.

Perdrizet II, G.A., Artsimovitch, I., Furman, R., Sosnick, T.R., et Pan, T. (2012). Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc Natl Acad Sci U S A* 109, 3323-3328.

Perez, J.C., Shin, D., Zwir, I., Latifi, T., Hadley, T.J., et Groisman, E.A. (2009). Evolution of a bacterial regulon controlling virulence and Mg^{2+} homeostasis. *PLoS Genet.* 5.

Perwez, T., et Kushner, S.R. (2006). RNase Z in *Escherichia coli* plays a significant role in mRNA decay. *Mol. Microbiol.* 60, 723-737.

Peters, J.M., Mooney, R.A., Grass, J.A., Jessen, E.D., Tran, F., et Landick, R. (2012). Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes Dev.* 26, 2621-2633.

Philippe, C., Eyermann, F., Bénard, L., Portier, C., Ehresmann, B., et Ehresmann, C. (1993). Ribosomal protein S15 from *Escherichia coli* modulates its own translation by trapping the ribosome on the mRNA initiation loading site. *Proc. Natl. Acad. Sci. U. S. A.* 90, 4394-4398.

Powell, B.S., Courtl, D.L., Nakamura, Y., Rivas, M.P., et Turnbough, C.L. (1994). Rapid confirmation of single integration by PCR copy lambda prophage. *Nucleic Acids Res.* 22, 5765-5766.

Prehna, G., Zhang, G., Gong, X., Duszyk, M., Okon, M., McIntosh, L.P., Weiner, J.H., et Strynadka, N.C.J. (2012). A protein export pathway involving *Escherichia coli* porins. *Structure* 20, 1154-1166.

Proshkin, S., Rahmouni, A.R., Mironov, A., et Nudler, E. (2010). Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation. *Science.* 328, 504-508.

Prost, L.R., Daley, M.E., Le Sage, V., Bader, M.W., Le Moual, H., Klevit, R.E., et Miller, S.I. (2007). Activation of the Bacterial Sensor Kinase PhoQ by Acidic pH. *Mol. Cell* 26, 165-174.

Pyle, A.M. (2002). Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem.* 7, 679-690.

Quinlan, A.R., et Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841-842.

Rabhi, M., Espéli, O., Schwartz, A., Cayrol, B., Rahmouni, A.R., Arluison, V., et Boudvillain, M. (2011). The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. *EMBO J.* 30, 2805-2816.

Rangan, P., et Woodson, S.A. (2003). Structural requirement for Mg^{2+} binding in the group I intron core. *J. Mol. Biol.* 329, 229-238.

Régnier, P., et Portier, C. (1986). Initiation, attenuation and RNase III processing of transcripts from the *Escherichia coli* operon encoding ribosomal protein S15 and polynucleotide phosphorylase. *J. Mol. Biol.* 187, 23-32.

Regonesi, M.E., Del Favero, M., Basilico, F., Briani, F., Benazzi, L., Tortora, P., Mauri, P., et Dehò, G. (2006). Analysis of the *Escherichia coli* RNA degradosome composition by a proteomic approach. *Biochimie* 88, 151-161.

Regulski, E.E., Moy, R.H., Weinberg, Z., Barrick, J.E., Yao, Z., Ruzzo, W.L., et Breaker, R.R. (2008). A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Mol. Microbiol.* 68, 918-932.

Reilman, E., Mars, R.A.T., Dijl, J.M. Van, et Denham, E.L. (2014). The multidrug ABC transporter BmrC / BmrD of *Bacillus subtilis* is regulated via a ribosome-mediated transcriptional attenuation mechanism. *Nucleic Acids Res.* 42, 11393-11407.

Repoila, F., et Gottesman, S. (2001). Signal Transduction Cascade for Regulation of RpoS :

Temperature Regulation of DsrA. *J. Bacteriol.* *183*, 4012-4023.

Roberts, J.W. (1969). Termination factor for RNA synthesis. *Nature* *224*, 1168-1174.

Rochat, T., Bouloc, P., et Repoila, F. (2013). Gene expression control by selective RNA processing and stabilization in bacteria. *FEMS Microbiol. Lett.* *344*, 104-113.

Ross, W., Vrentas, C.E., Sanchez-Vazquez, P., Gaal, T., et Gourse, R.L. (2013). The magic spot: A ppGpp binding site on *E. coli* RNA polymerase responsible for regulation of transcription initiation. *Mol. Cell* *50*, 420-429.

Roth, A., Winkler, W.C., Regulski, E.E., Lee, B.W.K., Lim, J., Jona, I., Barrick, J.E., Ritwik, A., Kim, J.N., Welz, R., et al. (2007). A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.* *14*, 308-317.

Ruff, E.F., Record, M.T., et Artsimovitch, I. (2015). Initial events in bacterial transcription initiation. *Biomolecules* *5*, 1035-1062.

Ruiz de los Mozos, I., Vergara-Irigaray, M., Segura, V., Villanueva, M., Bitarte, N., Saramago, M., Domingues, S., Arraiano, C.M., Fechter, P., Romby, P., et al. (2013). Base Pairing Interaction between 5'- and 3'-UTRs Controls *icaR* mRNA Translation in *Staphylococcus aureus*. *PLoS Genet.* *9*.

Santangelo, T.J., et Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase

runs stop sign (antitermination NusB). *Nat Rev Microbiol* 9, 319-329.

Schmidt, M.C., et Chamberlin, M.J. (1987). nusA Protein of *Escherichia coli* is an efficient transcription termination factor for certain terminator sites. *J. Mol. Biol.* 195, 809-818.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Holton, J.M., et Cate, J.H.D. (2005). Structures of the Bacterial Ribosome at 3.5 Å Resolution. *Science*. 827-835.

Selinger, D.W., Saxena, R.M., Cheung, K.J., Church, G.M., et Rosenow, C. (2003). Global RNA Half-Life Analysis in *Escherichia coli* Reveals Positional Patterns of Transcript Degradation. 216-223.

Sen, R., Qayyum, M.Z., Muteeb, G., Chalissery, J., et Vishalini, V. (2008). Nus Factors of *Escherichia coli*. *EcoSal Plus* 3, 1-16.

Shajani, Z., Sykes, M.T., et Williamson, J.R. (2011). Assembly of Bacterial Ribosomes. *Annu. Rev. Biochem.* 80, 501-526.

Shi, J., Jin, Y., Bian, T., Li, K., Sun, Z., Cheng, Z., Jin, S., et Wu, W. (2015). SuhB is a novel ribosome associated protein that regulates expression of MexXY by modulating ribosome calage in *Pseudomonas aeruginosa*. *Mol. Microbiol.* 98, 370-383.

Shin, J.H., Wakeman, C.A., Goodson, J.R., Rodionov, D.A., Freedman, B.G., Senger, R.S., et Winkler, W.C. (2014). Transport of Magnesium by a Bacterial Nramp-Related Gene. *PLoS*

Genet. 10.

Simons, R.W., Houman, F., et Kleckner, N. (1987). Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* 53, 85-96.

de Smit, M.H., et van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci.* 87, 7668-7672.

Smith, R.L., et Maguire, M.E. (1998). Microbial magnesium transport: Unusual transporters searching for identity. *Mol. Microbiol.* 28, 217-226.

Spencer, P.S., et Barral, J.M. (2012). Genetic Code Redundancy and Its Influence on the Encoded Polypeptides. *Comput. Struct. Biotechnol. J.* 1, e201204006.

Spickler, C., et Mackie, G.A. (2000). Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J. Bacteriol.* 182, 2422-2427.

Spinelli, S. V., Pontel, L.B., García Vescovi, E., et Soncini, F.C. (2008). Regulation of magnesium homeostasis in Salmonella: Mg²⁺ targets the mgtA transcript for degradation by RNase E. *FEMS Microbiol. Lett.* 280, 226-234.

Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Castillo, L.P., Ray, D., Van Bakel, H., Hughes, T.R., et Kushner, S.R. (2011). Analysis of *Escherichia coli* RNase e and RNase III

activity *in vivo* using tiling microarrays. *Nucleic Acids Res.* *39*, 3188-3203.

Steitz, T.A. (2008). A structural understanding of the dynamic ribosome machine. *Nat. Rev. Mol. Cell Biol.* *9*, 242-253.

Stringer, A.M., Currenti, S., Bonocora, R.P., Baranowski, C., Petrone, B.L., Palumbo, M.J., Reilly, A.A., Zhang, Z., Erill, I., et Wade, J.T. (2014). Genome-scale analyses of *Escherichia coli* and *Salmonella enterica* AraC reveal noncanonical targets and an expanded core regulon. *J. Bacteriol.* *196*, 660-671.

Sudarsan, N., Wickiser, K.J., Nakamura, S., Ebert, M.S., et Breaker, R.R. (2003). An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Genes Dev.* *17*, 2688-2697.

Sudarsan, N., Lee, E.R., Weinberg, Z., Moy, R.H., Kim, J.N., Link, K.H., et Breaker, R.R. (2008). Riboswitches in Eubacteria Sense the Second Messenger Cyclic Di-GMP. *Science*. *321*, 411–413.

Surratt, C.K., Milan, S.C., et Chamberlin, M.J. (1991). Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription. *Proc. Natl. Acad. Sci.* *88*, 7983-7987.

Takyar, S., Hickerson, R.P., et Noller, H.F. (2005). mRNA helicase activity of the ribosome. *Cell* *120*, 49-58.

Taylor, R.C., Webb Robertson, B.-J.M., Markillie, L.M., Serres, M.H., Linggi, B.E., Aldrich, J.T., Hill, E.A., Romine, M.F., Lipton, M.S., et Wiley, H.S. (2013). Changes in translational efficiency is a dominant regulatory mechanism in the environmental response of bacteria. *Integr. Biol.* 5, 1393.

Tock, M.R., Walsh, A.P., Carroll, G., et McDowall, K.J. (2000). The CafA protein required for the 5'-maturation of 16S rRNA is a 5'-end-dependent ribonuclease that has context-dependent broad sequence specificity. *J. Biol. Chem.* 275, 8726-8732.

Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., et Jung, K. (2013). Translation Elongation Factor EF-P Alleviates Ribosome Stalling at Polyproline Stretches. *339*, 82-86.

Urban, J.H., et Vogel, J. (2007). Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. *Nucleic Acids Res.* 35, 1018-1037.

Vanzo, N.F., Li, Y.S., Py, B., Blum, E., Higgins, C.F., Raynal, L.C., Krisch, H.M., et Carpousis, A.J. (1998). Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* 12, 2770-2781.

Vazquez-Laslop, N., Thum, C., et Mankin, A.S. (2008). Molecular Mechanism of Drug-Dependent Ribosome Stalling. *Mol. Cell* 30, 190-202.

Vogel, J., et Luisi, B.F. (2011). Hfq and its constellation of RNA. *Nat Rev Microbiol* 9,

578-589.

Voss, N.R., Gerstein, M., Steitz, T.A., et Moore, P.B. (2006). The Geometry of the Ribosomal Polypeptide Exit Tunnel. *J. Mol. Biol.* 360, 893-906.

Wahba, L., et Koshland, D. (2013). The Rs of biology: R-Loops and the regulation of regulators. *Mol. Cell* 50, 611-612.

Wang, J.X., et Breaker, R.R. (2008). Riboswitches that sense *S*-adenosylmethionine and *S*-adenosylhomocysteine. *Biochem. Cell Biol.* 86, 157-168.

Wang, H., Yin, X., Wu, M., Dambach, M., Curtis, R., et Storz, G. (2017). Increasing intracellular magnesium levels with the 31-amino acid MgtS protein. *Proc. Natl. Acad. Sci.* 114, 5689–5694.

Wang, J., Rennie, W., Liu, C., Carmack, C.S., Pr, K., Ding, Y., Caron, M., Mass, E., et Wade, J.T. (2015). Identification of bacterial sRNA regulatory targets using ribosome profiling. *Nucleic Acids Res.* 43, 10308-10320.

Wang, J.X., Lee, E.R., Rivera Morales, D., Lim, J., et Breaker, R.R. (2008). Riboswitches that Sense *S*-adenosylhomocysteine and Activate Genes Involved in Coenzyme Recycling. *Mol. Cell* 29, 691–702.

Waters, L.S., et Storz, G. (2009). Regulatory RNAs in Bacteria. *Cell* 136, 615-628.

Webb, M. (1966). The Utilization of Magnesium by Certain Gram-Positive and Gram-negative Bacteria. *J. Gen. Microbiol.* 43, 401-409.

Weinberg, Z., Barrick, J.E., Yao, Z., Roth, A., Kim, J.N., Gore, J., Wang, J.X., Lee, E.R., Block, K.F., Sudarsan, N., et al. (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res.* 35, 4809-4819.

Weixlbaumer, A., Leon, K., Landick, R., et Darst, S.A. (2013). Structural basis of transcriptional pausing in bacteria. *Cell* 152, 431-441.

Wilson, D.N., Arenz, S., et Beckmann, R. (2016). Translation regulation via nascent polypeptide-mediated ribosome stalling. *Curr. Opin. Struct. Biol.* 37, 123-133.

Winkler, W., Nahvi, A., et Breaker, R.R. (2002a). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419, 952-956.

Winkler, W.C., Cohen-Chalamish, S., et Breaker, R.R. (2002b). An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat. Struct. Mol. Biol.* 9, 15908–15913.

Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., et Breaker, R.R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428, 281-286.

Wolin, S.L., et Walter, P. (1988). Ribosome pausing and stacking during translation of a eukaryotic mRNA. *EMBO J.* 7, 3559-3569.

Woolhead, C.A., Johnson, A.E., et Bernstein, H.D. (2006). Translation Arrest Requires Two-Way Communication between a Nascent Polypeptide and the Ribosome. *Mol. Cell* 22, 587-598.

Woolstenhulme, C.J., Guydosh, N.R., Green, R., et Buskirk, A.R. (2015). High Precision Analysis of Translational Pausing by Ribosome Profiling in Bacteria Lacking EFP Christopher. *Cell Rep* 11, 13-21.

Wright, M.H., Adelskov, J., et Greene, A.C. (2017). Bacterial DNA Extraction Using Individual Enzymes and Phenol/Chloroform Separation †. *J. Microbiol. Biol. Educ.* 18, 1-3.

Yamamoto, K., Ogasawara, H., Fujita, N., Utsumi, R., et Ishihama, A. (2002). Novel mode of transcription regulation of divergently overlapping promoters by PhoP, the regulator of two-component system sensing external magnesium availability. *Mol. Microbiol.* 45, 423-438.

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature* 289, 751-758.

Yuzenkova, Y., Gamba, P., Herber, M., Attaiech, L., Shafeeq, S., Kuipers, O.P., Klumpp, S., Zenkin, N., et Veening, J.W. (2014). Control of transcription elongation by GreA determines rate of gene expression in *Streptococcus pneumoniae*. *Nucleic Acids Res.* 42, 10987-10999.

Zhang, N., Ren, X., Zhu, D., Li, D., et Wang, D. (2010). Crystallization and preliminary crystallographic studies of CorC, a magnesium-ion transporter. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 66, 681-683.

Zhao, G., Kong, W., Weatherspoon-Griffin, N., Clark-Curtiss, J., et Shi, Y. (2011). Mg²⁺ facilitates leader peptide translation to induce riboswitch-mediated transcription termination. *EMBO J.* 30, 1485-1496.

Zhou, K., Zhou, L., Lim, Q.E., Zou, R., Stephanopoulos, G., et Too, H. (2011). Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. *BMC Mol. Biol.* 12.

Zukher, I., Novikova, M., Tikhonov, A., Nesterchuk, M. V., Osterman, I.A., Djordjevic, M., Sergiev, P. V., Sharma, C.M., et Severinov, K. (2014). Ribosome-controlled transcription termination is essential for the production of antibiotic microcin C. *Nucleic Acids Res.* 42, 11891-11902.

Zurawski, G., Brown, K., Killingly, D., et Yanofsky, C. (1978). Nucleotide sequence of the leader region of the phenylalanine operon of *Escherichia coli*. *Proc. Natl. Acad. Sci.* 75, 4271-4275.

Zwir, I., Shin, D., Kato, a, Nishino, K., Latifi, T., Solomon, F., Hare, J.M., Huang, H., et Groisman, E. A (2005). Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*. *Proc Natl Acad Sci U S A* 102, 2862-2867.

